OCT 1 0 1931

The Journal of EDUCATIONAL PSYCHOLOGY

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

HAROLD RUGG, *Editor*

VOL. XXII, No. 6 SEPTEMBER, 1931 $5.00 per year

CONTENTS

OUR NEED OF SOME SCIENCE IN PLACE OF THE WORD "INTELLIGENCE." *C. Spearman* 401
AN EXPERIMENTAL STUDY OF THE INFLUENCE OF MOTION PICTURE FILMS ON BEHAVIOR. *Frank N. Freeman and Carolyn Hoefer*. 411
INTELLIGENCE, MOTIVATION, AND ACHIEVEMENT. *Austin H. Turney* 426
THE "OMISSION" AS A SPECIFIC DETERMINER IN THE TRUE-FALSE EXAMINATION. *C. C. Weidemann* 435
FACTORS INVOLVED IN CHILDREN'S FRIENDSHIPS. *Gladys Gardner Jenkins*. 440
A STUDY OF CLASSROOM BEHAVIOR. *Willard C. Olson* 449
REPLY TO PROFESSOR KELLEY. *Karl J. Holzinger*. 455
THE INFLUENCE OF JAZZ AND DIRGE MUSIC UPON SPEED AND ACCURACY OF TYPING. *Milton B. Jensen*. 458
A PROOF THAT THE POINT FROM WHICH THE SUM OF THE ABSOLUTE DEVIATIONS IS A MINIMUM IS THE MEDIAN. *Paul Horst* 463
RELATION BETWEEN USE OF DIFFERENT PARTS OF SPEECH IN WRITTEN COMPOSITION AND MENTAL ABILITY. *Charles P. Loomis and Anna May Moran* . . . . . 465.
FALSIFICATION OF AGE: A FACTOR IN CHILD GUIDANCE. *Anna Cohen and Nathan Altrowitz* 476
NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION 479
A CORRECTION. 480

Published Monthly Except June to August by

Warwick & York, Inc.,

10 East Centre Street Baltimore, Md.

Entered as Second Class matter November 15, 1921, at the Post Office at Baltimore, Md., under the Act of March 3, 1879; additional entry as second class matter at York, Pa.,

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY

ESTABLISHED 1910

Editor: HAROLD RUGG

Board of Associate Editors

J. CARLETON BELL — ARTHUR I. GATES — RUDOLF PINTNER — E. L. THORNDIKE
FRANK N. FREEMAN — V. A. C. HENMON — L. M. TERMAN — GOODWIN B. WATSON

Assistant to the Editor: FRANCES M. FOSTER

The Journal is published monthly, except June to August, by Warwick and York, Inc., Baltimore, Md., and York, Pa.

The price is $5.00 a year, payable in advance; foreign postage 30 cents extra. Single current issues are 70 cents each, and subscriptions for less than a year are charged at single issue rate for each number ordered. Back volumes are $6.00 each; single back numbers, 90 cents each, except, when more than three years old, the price is $1.00 each.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which the change is to take effect. Claims for non-receipt of an issue will not be honored unless made within two weeks after receipt of the next succeeding number.

Manuscripts for publication, books or other materials for review, and correspondence should be addressed to the publishers.

WARWICK AND YORK, INC., *Baltimore, Maryland*

TEACHERS' HEADQUARTERS

European Plan

Entirely Fire-Proof

Centrally Located

EDWARD DAVIS
Manager

HOTEL RENNERT
BALTIMORE, MARYLAND

The Psychology of Learning
1928 Revision — BY WILLIAM H. PYLE
$2.20 + 12¢ postage
Warwick & York, Inc., *Baltimore, Maryland*

The Acid Test

What more exacting test can there be of dictionary leadership than the judgment of the educational world?

Every State that has adopted an unabridged dictionary for use in its schools has chosen exclusively

WEBSTER'S NEW INTERNATIONAL DICTIONARY

The Colleges voted overwhelmingly in favor of Webster as the standard of pronunciation, in answer to questions submitted by the Chicago Woman's Club.

Nearly 100% of all schoolbooks are based on the New International for pronunciation, spelling, compounding, and division of words.

The New International has been universally accepted by those best fitted to judge.

It is in accord with the best modern practice.

Use it in your own schools.

Write for our free booklets of interesting ways to teach the use of the dictionary.

GET THE BEST

G. & C. MERRIAM COMPANY
SPRINGFIELD, MASS.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XXII | September, 1931 | Number 6

OUR NEED OF SOME SCIENCE IN PLACE OF THE WORD "INTELLIGENCE"

C. SPEARMAN

I. Contributions of Professor Dearborn

Not long ago the present writer ventured to criticize a book by Professor Dearborn on "Intelligence Tests."[1] He has now been good enough to respond;[2] and on one point at any rate I hasten to acknowledge a misunderstanding. I had complained that his book, whilst referring copiously to the small and controversial part of a work of mine,[3] had nevertheless taken no notice of the large and constructive part. In reply, he now states that his book, although published in 1928 (June) consisted really of some old lecture notes written two years earlier; hence, he says, its contents did not refer at all to the said work of mine published in 1927 (March), but instead to another work written by me as far back as 1922.[4] Although not without surprise at his manner of making books, I must admit that it has here rendered my criticism inapplicable; and so, with apologies, I withdraw it.

However Professor Dearborn, after thus making his own defence, passes over to a counter-attack. I will quote his own words, in order to render justice, not only to his purport, but also to his style. He charges me with "a brazen, if belated, effort to steal some of Binet's thunder," and supports this charge by alleging that "Binet did have some ideas of his own, preciously like the common factor of Spearman."

---

[1] Houghton, Mifflin and Co.
[2] This Journal, October, 1930.
[3] "The Abilities of Man," Macmillan, 1927.
[4] "Nature of Intelligence and Principles of Cognition," Macmillan, 1922.

To begin with, I should like to congratulate him on his choice of objective. Certain critics of our theory of Two Factors make a pretence of attacking it fundamentally when in truth they are only dealing with unimportant details; what is really essential in our theory they tacitly appropriate to themselves (mostly under new names). Not so Professor Dearborn. The point of attack chosen by him is both novel and vital. He would appear to charge us with nothing less than plagiarism; the very hub on which our theory rolls, the general factor itself, this he accredits to our earliest antagonist, Binet.

Despite this tribute to his honesty at any rate, in another respect Professor Dearborn, to speak frankly, is disappointing. Since he takes up a position of such grave moment, one might expect to find him entrenched behind a great mass of specific and cogent evidence. Actually, he brings forward no evidence whatever. Although my own representation of Binet's work had from end to end been accompanied by citations of that writer's very words, Prof. Dearborn in his turn does not add a single new quotation; nor does he even indicate any fresh way of interpreting those cited by me. Instead of any such thing, he gives us simply his personal conclusions. Such a procedure may be well enough in the class room, or even in electioneering hustings. But on the arena of science, just as in the courts of justice, I submit that it is too high-handed. We require the actual facts on which the assertions claim to be founded, so that we may have an opportunity to examine such claim for ourselves.

II. Dilemma of Binet

Attempting to make good this default on his part, let us first of all see what was the standpoint of Binet *before* the appearance of "'General Intelligence,' objectively determined and measured,"[1] which was published by the present writer in April 1904, and already contained all the essentials of the theory of Two Factors. In particular, its main object—as shown in the very title—was to introduce a "general intelligence" that admitted of measurement by tests. Before this publication, the standpoint of Binet seems to have been most characteristically expressed as follows:

> We propose the study of the ten following processes. Memory, the nature of mental images, imagination, attention, the faculty of understanding, suggestibility, the aesthetic sentiment, moral sentiment, muscular force, and force of will, motor ability, and coup d'oeil. These are, we believe, mental faculties that differ

[1] *American Journal of Psychology*, Vol. XV.

greatly from one individual to another, and are such that the knowledge of their state with an individual gives us a general idea of that individual.[1]

He then proceeds to supply us with several tests for each faculty. Memory is to be tested by geometric designs, phrases, music, and colours. Mental images, by the nature of the errors of reproduction. Imagination, by ink blots, composition of pictures, and combination of words into phrases. Attention, by regularity of reaction-times, and by success in counting two metronomes simultaneously. Comprehensions, by definitions, by perception of resemblances and differences of meaning, and by criticism of phrases.[2] Of any "general factor" he does not so far seem to have manifested any trace. Indeed, even the term "intelligence" does not appear to be mentioned. And to the best of my knowledge, he made no radical departure from this standpoint right up to our aforesaid publication of the theory of Two Factors in 1904.

But he certainly did so two years *after* that date. For then he and Simon brought forward their world-famous "Metric Scale of Intelligence." They wrote:

> Our aim is in no way to study, to analyse and disengage the aptitudes of those who are inferior in intelligence. That will be the object of a later work. Here, we will restrict ourselves to appreciating and measuring their *intelligence in general;* we will fix their intellectual level; and to give an idea of that level, we will compare it with that of normal children of the same age, or of analogous level.[3]
>
> But here, we must understand one another on the meaning of that word so vague and so comprehensive: "intelligence." Almost all the phenomena with which psychology is concerned are phenomena of intelligence; a sensation, a preception, are intellectual manifestations as much as is a reasoning. Ought we then to introduce into our examination the measurement of sensation, as is done by psycho-physicists? Ought we to put into our tests the whole of psychology?
>
> A little reflection has shown us that that would be a great loss of time. There is *within intelligence*, so it seems to us, a *fundamental organ*, that whose lack or perversion is of the greatest importance for practical life. This is judgment, otherwise called good sense, practical sense, initiative, the faculty of adapting oneself. To judge well, to understand well, to reason well, these are the essential springs of intelligence.[4]

These words, at first reading, were delightful enough. For between this measurable "intelligence in general" thus newly advocated by

---

[1] *Année Psychol.*, Vol. II, published in 1896. The date on the volume is 1895; but this, it appears, always refers to the *previous* year (supposely the time that the actual research was effected).

[2] *Ibidem.*

[3] *Année Psychol.*, Vol. XI, published in 1906, see above:

[4] *Ibidem,*

Binet and the measurable "general intelligence" introduced two years earlier by myself, there seemed to be no difference. And as for his "fundamental organ within the intelligence," this appeared to be only a special interpretation of our factor "*g*." There are accordingly some grounds for concurring with Prof. Dearborn to the extent that *by this time* (1906) Binet did have the concept of a general factor in intelligence.

But of any such likeness between his view and ours, Binet himself seems not to say a word. As I pushed on with the perusal of his work, expecting every moment to come upon his expression of cordial agreement, I encountered indeed a review of our theory of Two Factors; but this only announced his disapproval of it, on the ground of it being "stuffed with mathematics"!

The fact seems to be that, so far as theorizing is concerned, his adoption of our general factor was only half-hearted. He was torn in two opposite directions. On the one hand, he naturally endeavoured bring his utterances into harmony with his new metric scale. But on the other hand, he continually regressed back to the old faculty doctrine; for upon this had been built up, and now stood irrevocably founded, his whole general psychological outlook. Even in this very metric scale that we have been discussing, he flagrantly lapses into the facultism again; for he expressly assigns to each of his tests its own special faculty or faculties. Thus his test of cutting paper is said by him to be one of "voluntary attention, reasoning, and visual imagination." His discrimination between objects is called a test of "ideation, the notion of difference, and to some extent the spirit of observation." His repetition of digits is said to involve the faculties of "immediate memory and of voluntary attention." His test of pictures—to which he subsequently attributed such dominant importance—is here only taken as "destined to seek whether there exist associations between images and their names." In the second version of the metric scale published three years later, all such indications of special faculties to be measured by the respective tests was, indeed, abandoned. But very soon afterwards, they were re-introduced and more explicitly than ever. For now he goes so far as to write that:

> The mental faculties of each subject are independent and unequal . . . Our mental tests are always specific in their scope; each suits the analysis of a single faculty.[1]

[1] Les Idées modernes sur les Enfants, 1909.

In these words, the general factor or "fundamental organ" would seem to have quite faded out of the picture.

But whilst thus in theoretical discussion ever regressing back to his ingrained facultism in the actual practice of measuring the logic of facts drove him irresistibly more and more along the new path of Two Factors. Already in 1906, as mentioned, he had indicated his intention to measure the "level" of his "intelligence in general." But how was such a feat possible? The word "general" means something common to all or nearly all of a class, as exemplified in the phrases a "general direction" or the "general government." But with the faculties, he himself stated (as we have just seen) that no such common level exists. His sole resource was to take "general" not in the sense of common, but in that of middle or average. Such averaging, however, was just what we had advocated already in 1904. Brilliantly original, then, as was his procedure in detail, in essence it was following us.

III. Limitations to the Validity of Averaging

But was he here really entitled to follow us? The procedure of averaging, wide-spread and invaluable as it is in both physical and mental science, has nevertheless its limitations. It is only legitimate when certain postulates are fulfilled. One of these is the unequivocality of the domain at issue. You cannot, for instance, obtain the average height of an American without first settling who is to count as one. And then, having got the domain definitely settled, it is imperative to take the whole of this into the reckoning. Thus, you cannot claim to have got the monthly average of your expenditure during any year if you leave out of calculation, say, December. A third postulate is that your reckoning should present no overlaps; thus, it would not do to bring in December several times. Still more fundamental is the need of comparable units. You may indeed equally well derive your average expenditure by taking every day, or every week, or every month. But you would not be rational if you mixed up days, weeks, and months indiscriminately together.

Now, on the theory of faculties, one at least of these postulates appears impossible to satisfy; namely, that of finding any common unit by which the different faculties can be brought into mutual connection. Thus, how shall a common unit be found to make memory comparable with judgment? We might take the whole of memory as one single unit whilst we divided up judgment into several kinds each constituting a unit of its own; or else equally well we might

take judgment as being one and subdivide the memory. The one basis for averaging would favor some individuals, and the other basis others. Thus the score reached by averaging becomes arbitrary, therefore invalid.

And the matter becomes much worse still, if we pass from the clear but untenable theory of faculties to the extremely vague theory of "intelligence" commonly favored. For here, all four of the said postulates fail to be satisfied. To find any common unit remains as impossible as before. In addition, there seems no way of avoiding large omissions; for it would appear hopeless to attempt to include in the tests every sort and description of "intelligence." Furthermore, there appear to be no means of arranging that none of the tests overlap each other; indeed, they often admittedly do so. And lastly, no agreement seems to be feasible as to how the domain of this "intelligence" should be delimited; for some authorities want to include one thing and others others.

Nor would the preceding difficulties be in any degree escaped by shifting from the notion of average to that of a "totality." For the latter also would involve every one of the aforesaid postulates (even—as might perhaps not be expected—that of a common unit). Nor yet would any relief be brought by substituting, as many later writers have done, the notion of a "sample." For this too would involve the same four postulates—with other troubles superposed.[1] To all this may be added that actually mental testers, far from having overcome these difficulties, have hitherto not so much as made any serious attempt at doing so.

But what is sauce for the goose, it may be urged, is also sauce for the gander. If all this averaging is so irrational when employed by the facultists and the intelligentists, how can it pretend to be satisfactory in the hands of the two-factorists? The reply is that in the latter case all the preceding difficulties do surprisingly dissolve away. For in it, no claim is raised that the result of combining together different tests is to constitute any rational average, or central value, or totality, or sample. The sole claim raised is that, in such a procedure, the specific factors tend to cancel out, leaving the general factor alone dominant. As was said originally:

> "The specific element can to a great extent be readily eliminated by varying and combining the kinds of test," whereby the general factor or "central function" stands out in corresponding exactness.[2]

[1] "The Abilities of Man," Chap. V.

[2] *American Journal Psychology*, Vol. XV, April, 1904.

For *this* purpose of cancelling out, omissions or overlappings—unless they are highly systematic—do not greatly interfere. Nor does even the manner in which the units are constituted. And as for the domain from which the tests are to be derived, this in the theory of two factors is no longer equivocal. Its domain extends to just as many abilities as may be observed to satisfy the two-factor criterion.

IV. Plea for an Awakening

What is the upshot of all this? Not unnaturally perhaps, Prof. Dearborn comes to the conclusion that, in spite of all, the situation is well enough; that, even if the current procedure of measuring the "mental ages" and the "IQ's" is indeed equipped with a wrong theory, it nevertheless admittedly is right in practice, which is what alone really matters.

Now, it is just against this attitude of quietism that I would here desire to make a lively protest. For in the first place, even the practice current in testing is only right with very crude approximation. It only succeeds ever in yielding a definite measurement at all if and when the sub-tests entering into the whole team of them are sufficiently numerous and *selected at random.* Now some sort of randomness does derive from the mere fact that these subtests are gathered together from highways and byways without any system or plan. But this facile mix-up is insufficient to secure that perfect randomness—in the sense of freedom from all bias—which statisticians mean by the term and which is needed to produce the definite "*g*." In order to obtain such freedom from bias, the procedure should not be without all plan, but on the contrary planned with great care. Whereas the present test-makers, quiet light-heartedly allow themselves to be strongly biassed towards several objectives, such as ease of application, self-consistency, correlations with other things, and so forth.

But there is a still worse objection to the current procedure. Even if this did happen to produce pure *g* (or any other definite measure), nevertheless when divorced from the theory out of which it sprang it would pay the penalty of forfeiting all significance; the mere average of sub-tests picked up and put together without rhyme or reason must needs present—at any rate until proof is given otherwise—the very acme of *meaninglessness.* And as for the customary dignifying of this gallimaufrey of tests by the title of "intelligence"—without so much as an attempt really to construct them in accordance with

any explicit principle—this device comes unpleasantly near to psychological characteristics.

Having regard to these two defects, then, those of indefiniteness and of meaninglessness, there is but small wonder that the current procedure from time to time commits, even in practice, blunders of great magnitude; so that there is only too good reason to suspect that many a child branded by test has thereby unjustly had his career wrecked. And this terrible fallibility of the tests is only rendered the more dangerous by the vast edifice of statistical data and calculations which are often ostentatiously piled upon them and which do but convert them into whited sepulchres.

But by what manner of means, it may be asked, is all this evil going to be made any the better by bringing on the scene the theory of Two Factors? Are not these also repeatedly charged with being entities of an obscure, speculative, or hypothetical kind? To say this about the factors is totally to misapprehend them, as we will now endeavour to show. These two factors, *g* and *s*, do not in themselves signify any entities whatever; they are nothing more than parts of actually observed measurements of ability. And these part-values have at least the virtue of perfect quantitative definiteness; that is to say they are, when taken rightly, "unique." But to acquire scientific meaning, a further step is required; there must be proof supplied that these part-values, besides being definite quantitatively in themselves, also serve as measures of some qualitatively definite psychological functions. And this, too, we may claim, has actually been done. In particular, the *g* has been found to measure "noegenesis."[1] This term denotes the following three actually and precisely observable mental functions. The first is the knowing of our own experience: we not only feel, but know that we feel; we not only strive, but know that we strive; we not only know, but know that we know. The second of the mental functions consists in that, when any two or more items are perceived or thought of, we may cognize relation between them; for example, we may become aware that one thing is like another, or is evidence for another, or is the cause of it. An outstanding instance of this among tests is that of "synonyms and antonyms," where two words are given and the testee has to decide whether they are very alike or very unlike. Thirdly and lastly, when

[1] The "noe-" comes from the Greek νοῦς, and indicates that processes of this form alone produce knowledge. The "-genesis" indicate that such processes alone generate new cognitive content.

we perceive or think of any item and also of any appropriate relation, we may evoke the idea of the item which stands to the given one in the given relation. A well know instance among tests is that of "opposites," where the testee is given some word and is told to say what is the opposite to it. These second and third functions are usually characterised as "eductive."

Whilst thus urging that the noegenetic processes are conceived in a manner free from all indefiniteness, we may take the opportunity to defend them from the charge sometimes raised that they are only of a "logical" nature, not of a "psychological" one. What is here meant by logical has been left by these critics regrettably obscure. But possibly the notion is that these kinds of process have little or no importance for understanding, anticipating, and controlling the course of mental life. In reply, we must claim just the contrary. These three processes follow in general a settled sequence, and in such fashion that each prepares the way for the following one. Thus, the awareness of one's own experience supplies all the fundamental items of knowledge; next, between these items (possibly in some cases simultaneously with them) the relations are cognised; and finally the relations thus furnished are displaced to other items and thus produce all correlative and novel items which are the mainstay of mental creation and invention, including that super-invention which is behavior.[1] To recognise this sequence of the three steps and to explore the conditions which facilitate each of them should constitute one of the main objectives of all functional psychology, including in particular its applications to education, medicine, and industry. And all this is called "not psychological but logical"!

More subtle and therefore still more dangerous than any such direct charges against the noegenetic analysis is the not uncommon practice of tacitly ranking it with mere unevidenced assertions on the matter. For example, the view that mental tests involve noegenesis is put on the same footing with the assertion that they involve adaptibility or educability, and so forth. This is done regardless of the fact that the noegenetic view is based on the most elaborate and exhaustive measurements in all detail, whereas the other views have nothing behind them but unsupported statement.

But even this establishment of *g* as being a measure of the ability called noegenesis is only the beginning, by no means the sum, of the

[1] For a full but simple account of such displacement of relations, reference must be made to the present writer's "Creative Imagination," Appleton, 1930.

definite and scientifically significant contributions of the theory of Two Factors. In support of this claim, appeal may be made to the second part of "The Abilities of Man" quoted on the first page of this paper. For this whole part is packed, not with anything obscure or speculative, but with fundamental facts actually observed by a multitude of indefatigable workers during the past quarter of a century. And to all this light which the Two Factors have shed upon the working of the mind must be added that which they have thrown upon the current tests of the socalled general intelligence. For the *g* appears to constitute just that which is common to different tests or sets of tests, and thus to furnish the sole stable element in intelligence-testing. To this stable element, the tests at present in use supply only crude approximations; what they contain over and above this *g* varies in an incalculable and therefore unscientific manner from one such test to another.

There is yet another objection commonly raised to the theory of two factors. Here, the claim is advanced that this mode of analysing the scores made in a test is only permissive, not obligatory; other manners of analysis are proposed—or more often vaguely suggested as possible—to take its place. Against this attitude we must plead that, on the contrary, there are places for any number of different analyses; any two different ones may quite well be *both of them* true and even both of them valuable, or both of them the reverse. Each has to stand or fall independently; its fate must be settled solely by the amount of scientific information which it produces. For example, the definitely observed fact that *g* measures noegenesis remains quite unaffected by what any other factors arising from any different analysis may or may not happen to measure. If any factors other than *g* and *s* do also happen to render useful services, then these services can only supplement one another, and thus enhance each other's values.

All in all, then, the plea is here raised that the current procedure of testing "intelligence" needs to be aroused from its self-complacent slumber. The present real basis of it is only a practice that has been borrowed from another theory. And this practice fails to be satisfactory because, when the practice was borrowed, the theory itself was left behind. By so doing, a sacrifice was made of precision; of meaningfulness; and above all, of an immense amount of observed fundafental facts. This disastrous situation has been largely masked, but at the same time really aggravated, by usurping the pretentious and wellnigh fraudulent title of "intelligence."

AN EXPERIMENTAL STUDY OF THE INFLUENCE OF MOTION PICTURE FILMS ON BEHAVIOR[1]

FRANK N. FREEMAN

University of Chicago

AND

CAROLYN HOEFER

Elizabeth McCormick Memorial Fund

The purpose of this study was to determine whether motion picture films add to the effectiveness of instruction which is designed to incite children to perform specific acts. The problem is one phase of the more general problem of the comparative effect of motion pictures and of other influences in guiding or modifying conduct. In this particular instance the instruction, of which the motion pictures are a part, was designed to influence conduct largely through giving clear-cut information and pointing out the applications of this information to conduct. The motion pictures which were used should have been chiefly of an informational nature to harmonize with the character of the general unit of instruction. Suitable informational films could not be obtained, however, and the films which were used were cast in story form. They were designed to appeal to the emotions as well as to the intellect.

The instruction dealt with the care of the teeth. This subject was chosen because it was believed that any change in the children's behavior as a result of instruction could be more objectively measured in this field than in any other which was open to investigation. The Elizabeth McCormick Memorial Fund, moreover, was in a position to provide systematic curriculum material on this subject and to make contacts with officials of the Health Department of Chicago and of the Public Schools as a result of the co-operative study of other problems in health education.

---

[1] This study was made possible by a grant from the Committee for the Study of the Social Values of Motion Pictures of the Payne Fund. Grateful acknowledgment is also made to Dr. George Wandel of the American Dental Association and Dr. Lon Morrey of the Health Department of Chicago for the loan of the films and for assistance in the preparation of dental lessons, to the Chicago Department of Health and the Superintendent of Schools for the assistance given in the dental examinations.

The experiment was made in two public schools of Chicago. Two fifth and two sixth grades in one school were taught with the aid of the films, while two fifth and two sixth grades in another school in a comparable neighborhood were taught without the films. In order that the instruction might be comparable, except for the films, it was all given by the same person, Miss Mildred Dawson. Miss Dawson also served as general assistant in the experiment. She is a very well-trained and competent teacher and supervisor.

The instruction was given in thirteen periods, from May 8 to May 22 inclusive. The following is the outline of topics:

Lesson 1.—Know your teeth.
Shape and function of teeth.
Lesson 2.—Structure of a tooth.
Lesson 3.—Diet for teeth.
Lesson 4.—Diet for teeth.
Lesson 5.—Deciduous teeth.
Lesson 6.—Permanent teeth.
Lesson 7.—Six-year molars.
Lesson 8.—Film or story.
Lesson 9.—Malocclusion and mastication.
Lesson 10.—Causes of decay
Lesson 11.—Prevention of decay, and review.
Lesson 12.—Care of the teeth, and review.
Lesson 13 —Film or writing slogans.

Charts and diagrams were used in the instruction of both groups to supplement oral instruction. The time occupied by showing the film to the experimental group was employed in reading a story and in conducting an exercise in writing slogans in the control group.

The two films which were shown to the experimental group are entitled Tommy Tucker's Tooth and Clara Cleans Her Teeth. The first tells the story of a boy who was refused a job for which he applied because of his unsightly teeth, but who later had his teeth looked after and was then accepted. The second tells of a girl who enters a new school where the children have been trained to care for their teeth, and who is ostracized because of the poor condition of her teeth and her indifference to it. She, too, reforms and becomes socially acceptable to her schoolmates. Some information is given in the films, but it is subordinate to the stories.

The effectiveness of the instruction in the experimental and control groups was measured in four ways, namely:

1. An information test, given at the beginning and the end of the period of instruction;
2. A questionnaire in which the children reported on the care of their teeth at home and the food which they ate;
3. The material brought in by the children, consisting of clippings of various sorts, original material written by the children, booklets read, literature taken home, and score cards which the children filled out;
4. Condition of the teeth as determined by a dental examination given at the beginning and at the end of the experiment.

Results. The Information Test

The information test was given at the beginning and the end of the experiment to both groups. It consists of thirty-six multiple-choice questions. The questions relate to the structure and the development of the teeth, the causes of defective teeth, and the care of the teeth. Through difficulty in tabulation, four of the questions were omitted, leaving thirty-two. The information test is not a direct measure of behavior. Since the purpose of much of the instruction, however, was to give information, it is important to know whether the films had any influence directly or indirectly upon the amount of information which the children obtained. A comparison of the gains in information in the two groups is, therefore, presented.

The differences in the average ability of the children in the individual grades of the two schools was so great that it was thought necessary to make matched groups based on scores on the intelligence tests and the initial scores on the information test. The matched groups were formed, not by matching individuals, but by selecting individuals from the two groups so that there would be an approximately equal number in each class of the distribution. This form of matching gives not only equal average scores, but also an approximately equal number at the various levels of ability.

Table I gives the data on the results of the information test from the equalized groups. It will be observed that the average intelligence quotient and the average initial score of the two groups are practically identical.

It is apparent that both the boys and girls of the control group made somewhat higher average scores on the information test than did the children of the experimental group, but the difference is hardly statistically significant. We are safe in saying, however, that the films had no appreciable influence on the acquisition of information. This is, perhaps, not surprising, since the films were not designed primarily to

give information. They were designed rather to make the children interested in caring for their teeth. All we can say, then, is that these films, designed to awaken interest in the care of the teeth, did not stimulate the children who saw them to acquire more information than the children who did not see them. Interest in acquiring information about the teeth is evidently somewhat independent of interest in the care of the teeth, assuming that the films have stimulated interest in the care of the teeth. It is, of course, possible that they did not stimulate such interest. The inference that the two types of interest are independent, therefore, is not conclusive. It is contingent upon an

TABLE I.—THE SCORES OF THE EQUALIZED EXPERIMENTAL AND CONTROL GROUPS ON THE INFORMATION TEST AT THE BEGINNING AND THE END OF THE EXPERIMENT

Equalized groups[1]	No.	Average IQ	SD	Average Test I	SD	Average Test II	SD
Experimental school.							
Boys	55	100.98	15.247	14.91	3.561	27.73	5.992
Girls	53	101.32	11.474	15.40	4.003	28.62	4.180
Boys and girls	108	101.15	13.517	15.15	3.793	28.17	5.193
Control school.							
Boys	50	100.78	16.364	14.90	3.689	29.64	5.768
Girls	49	101.37	11.642	15.41	3.523	29.06	4.895
Boys and girls	99	101.07	14.235	15.15	3.622	29.35	5.370

[1] Equalized on basis of sex, intelligence rating and initial score on information test.

assumption for which we have as yet no evidence. We shall have to hold the conclusion in abeyance, therefore, until we have examined the evidence concerning the effect of the films in stimulating interest in the care of the teeth.

In order to discover whether the film was of special aid in giving particular items of information, the individual questions of the information test, which bore upon the activities involved in the care of the teeth, were tabulated separately. In only two questions does the experimental group give any clear evidence of being superior to the control group in the gains on the individual questions. The correct answer to the first of these questions is: The motion of the brush should be up and down; and to the second: To keep the teeth in good condition we should eat some food that requires much chewing. The method of

using the toothbrush was shown in both films, and the use of food which requires chewing was shown in the film, Tommy Tuckers' Tooth. The superiority of the film group in these two questions, therefore, is plausibly explained by the fact that the items of information or of activity to which they refer are specifically emphasized in the film.

The Results of the Questionnaire

At the end of the experiment, the children of both groups were given a score card to fill out. On this score card they were asked to report whether or not (1) they possessed a tooth-brush, (2) they brushed their teeth daily, (3) they used a dentifrice, (4) they ate vegetables daily, (5) they ate fruit daily, and (6) they drank milk daily.

The results of the questionnaire are tabulated in Table II. The table shows the percentage of children of each group who answer the

Table II.—The Responses to the Questionnaire on the Care of the Teeth and on Eating Given at the End of the Experiment

		Home care						Food					
	No.	Possession of tooth-brush		Teeth brushed daily		Use of dentifrice		Vege-tables eaten daily		Fruit eaten daily		Milk drunk daily	
		No.	Per cent	No.	Per cent	No.	Per cent	No.	Per cent	No.	Per cent	No.	Per cent
Experimental school.[1]													
Boys	92	92	100.00	90	97.83	88	95.65	85	92.39	88	95.65	91	98.91
Girls	100	99	99.00	98	98.00	95	95.00	94	94.00	91	91.00	95	95.00
Boys and girls	192	191	99.48	188	97.92	183	95.31	179	93.23	179	93.23	186	96.88
Control school.[2]													
Boys	89	86	96.63	66	74.16	83	93.26	81	91.01	87	97.75	89	100.00
Girls	81	80	98.77	72	88.89	77	95.06	74	91.36	78	96.30	77	95.06
Boys and girls	170	166	97.65	138	81.18	160	94.12	155	91.18	165	97.08	166	97.65

[1] Records of ten boys and six girls incomplete due to absence.
[2] Records of eight boys and four girls incomplete due to absence.

various questions in the affirmative. It is obvious that, if the children's testimony is to be relied upon, a large percentage of both groups perform faithfully the act which they were taught to perform in the instruction. Whether or not they would have given an equally good account of themselves at the beginning of the experiment we do not know. On account of the high scores in both groups and the uncertainty concerning the accuracy of the children's replies, small differ-

ences between the two groups are of no significance. The only item on which the difference may be significant is that concerning the daily brushing of teeth. Ninety-eight per cent of the children of the experimental group reported that they brushed their teeth daily whereas only eighty-one per cent of the control group made a similar report. Since the films emphasize the daily brushing of teeth very strongly, this difference may be significant. The films also emphasize the other items, however, and it is not altogether clear why there should be so much more difference in this item than in the others. So far as it goes, the data suggest that when a particular action is strongly emphasized in a film the emphasis may be effective. However, it is not at all clear why the control group should have fallen so much lower in brushing their teeth daily than in using dentifrice, or in eating the various kinds of food which were advised. Taken altogether, then, the results of the questionnaire do not indicate much one way or the other concerning the relative influence of the motion picture and of oral instruction.

Measures of Interest

As measures of interest a record was kept of certain specific voluntary activities on the part of the children. This record included, first, a record of all the clippings or original articles, pictures, poems, stories, etc., which were contributed by the children; second, the number of booklets which were read by the children of their own accord, the amount of literature which was taken home and the number of score cards concerning their reading which they voluntarily kept. The booklets, literature, and score cards were placed on the desk where the children could get them and they were told that they were free to use them if they wished to do so.

A summary tabulation of the contributions and voluntary acts is found in Table III. A comparison of the integral numbers in the table is qualified somewhat by the difference in the total number of pupils in the two groups. The number represented in the experimental group is 208 and in the control group 182. Some of the differences, however, are of such magnitude that this qualification is of minor importance.

It will be seen that the pupils of the experimental group brought in a much larger number of articles and pictures and contributed a larger number of original articles, cartoons, and pictures. The pupils of the control group brought in a somewhat larger number of poems

and drew a somewhat larger number of posters. The differences, however, are not so large as those in favor of the experimental group. Taken as a whole, it is clear that the pupils of the experimental group

TABLE III.—A RECORD OF CERTAIN VOLUNTARY ACTIVITIES CARRIED ON, USED AS A MEASURE OF INTEREST

	Experimental school (102 boys, 106 girls, 208 total) No.	Control school (97 boys, 85 girls, 182 total) No.
Contributions.		
Clippings.		
Article	53	7
Cartoon	3	5
Chart	7	1
Magazine	1	3
Model	0	1
Pamphlet	2	0
Picture	410	72
Poem	2	29
Poster	0	4
Story	2	1
Total	480	123
Average per child	2.31	0.68
Original.		
Article	16	3
Cartoon	25	4
Chart	1	4
Picture	40	9
Poem	10	4
Poster	0	14
Story	7	1
Total	99	39
Average per child	0.48	0.21
Grand Total	579	162
Average per child	2.78	0.89
Average booklets read	0.69	0.66
Average pieces of literature taken home	0.74	0.82
Average score cards filled out	0.46	0.35

contributed a much larger number of articles and of original materials of various sorts as reckoned in terms of the average number of contributions per child. In the number of booklets read, the literature taken home, and the score cards filled out, however, there is no significant

difference between the two groups. The balance is slightly in favor of the experimental group and the influence of the film appears to be somewhat significant. It is probably significant that the item in which the experimental group is most superior is in the number of pictures which were contributed. The sight of the motion picture apparently suggested to the children that they find similar pictures and bring them to the class. The results do not suggest a great difference in general interest or motivation between the two groups.

Change in the Condition of the Teeth

It was thought that a change in the condition of the teeth might be a more objective measure of the care which was exercised by the children or their parents and might, therefore, be the most objective measure of the effectiveness of instruction. The two types of changes which seem to be most significant are the change in the number of carious teeth and the change in the number of teeth which have received dental care. The reports of the dental examinations at the beginning and the end of the period of instruction were therefore examined to obtain a comparative measure of these two types of changes in the experimental and the control groups.

It is necessary to confine the comparison to the permanent teeth. There is a difference in practice among dentists in the treatment of deciduous teeth—some fill them, others do not.

Again, some deciduous teeth disappear between the first and second examination and the number of carious teeth would, therefore, be decreased without any effort on the part of the child or his parents.

The first permanent molars, sometimes referred to as six-year molars, were selected for examination because these are the permanent teeth which have been in the mouth the longest time. They, therefore, give the greatest opportunity for decay and for the exercise of dental care.

In interpreting the condition of the teeth or the change in the condition of the teeth it is necessary to take careful account of age. The eruption of these molars themselves is highly correlated with age and the eruption of the neighboring teeth, which have an influence upon decay, is also correlated with age. For this reason it was necessary to equalize the groups in accordance with the distribution of ages, in the same manner as was done for intelligence and initial score in the tabulation of the results of the information test. Since the teeth of girls erupt earlier on the average than those of boys, it is

necessary also to keep the results from the two sexes separate in the tabulation.

The first tabulation of a change in the condition of the teeth is given in Table IV. This table shows the increase or decrease in the number of carious teeth from the first to the second examination. The groups have been equalized with respect to chronological age, but obviously they have not been equalized with respect to the condition of the teeth at the first examination. In the experimental group there were 253 carious molars at the first examination whereas in the control group there were 318. This may indicate greater suscep-

TABLE IV.—CHANGE IN THE NUMBER OF CARIOUS FIRST PERMANENT MOLARS FROM THE FIRST TO THE SECOND EXAMINATION AFTER THE GROUPS HAVE BEEN EQUALIZED FOR AGE

	No.	Average CA months	First examination		Second examination	
			No. carious molars	Average per child	No. carious molars	Average per child
Experimental school.						
Boys	86	138.37	122	1.42	137	1.59
Girls	81	135.64	121	1.49	126	1.56
Boys and girls	167	137.04	243	1.46	263	1.57
Control school.						
Boys	86	137.70	168	1.95	167	1.94
Girls	81	135.56	150	1.85	133	1.64
Boys and girls	167	136.68	318	1.90	300	1.80

tibility of the children of the control group to caries or it may indicate that these children have had less care. The average number of carious molars per child increased in the experimental group and decreased in the control group. At first sight this may appear to indicate that the instruction was more efficient in the control group than in the experimental group. Account must be taken, however, of the difference in the condition of the teeth of the two groups at the beginning of the experiment. It is at least possible that the greater increase in the carious teeth of the children in the experimental school is due to the fact that there was greater opportunity for increase.

In order to furnish a check on this possibility, the groups were equalized not only with respect to chronological age, but also with respect to the condition of the first permanent molars at the first examination. In other words, the children were selected in the two groups so as to provide an equal number of carious molars at first examination. The comparison is shown in Table V. The difference

TABLE V.—CHANGE IN THE NUMBER OF CARIOUS FIRST PERMANENT MOLARS FROM THE FIRST TO THE SECOND EXAMINATION AFTER THE GROUPS HAVE BEEN EQUALIZED FOR INITIAL CONDITION OF THESE MOLARS

	No.	Average CA	First examination		Second examination	
			No. of first permanent molars carious	No. carious per child	No. of first permanent molars carious	No. carious per child
Experimental school.						
Boys	47	139.02	92	1.96	79	1.68
Girls	45	133.20	95	2.11	86	1.91
Boys and girls	92	136.17	187	2.03	165[1]	1.79
Control school.						
Boys	47	138.94	91	1.94	85	1.81
Girls	45	133.82	95	2.11	86	1.91
Boys and girls	92	136.43	186	2.02	171[2]	1.87

[1] Twenty-two molars carious at first examination and in "perfect condition" at second examination.

[2] Fifteen molars carious at first examination and in "perfect condition" at second examination.

in the two groups is now reversed. The number of carious teeth per child decreases in both groups and slightly more for the experimental than for the control group. Apparently the unfavorable showing of the experimental group in comparison to the control in the earlier table was due to a difference in the initial condition of the teeth. The difference in favor of the experimental group in the equalized groups is hardly great enough to have statistical significance, particularly as it is all due to a difference in the case of the boys. It may be that we should expect no difference. Even the best of care could probably affect but little change in the period of the experiment and the difference due to a difference in the effectiveness of the instruction in the two groups might be so minute as to be imperceptible when

TABLE VI.—NUMBER OF FIRST PERMANENT MOLARS WHICH WERE IN "PERFECT CONDITION" AT THE FIRST EXAMINATION AND WHICH WERE CARIOUS FILLED OR EXTRACTED AT THE SECOND EXAMINATION

	Experimental school										Control school									
	No.	No.	Av. CA	Av. IQ	Carious		Filled		Extracted		Received dental care		No.	Av. CA	Av. IQ	Carious		Filled		Extracted		Received dental care	
					No.	Per cent	No.	Per cent	No.	Per cent	No.	Per cent				No.	Per cent	No.	Per cent	No.	Per cent	No.	Per cent
No. 3 Not Carious First Examination																							
Boys	86	27	137.63	101.00	6	22.2	1	3.7	1	3.7	2	25.0	19	138.47	101.11	3	15.8	1	5.3	1	5.3	2	40.0
Girls	81	28	136.04	103.93	5	17.9	4	14.3	0	0.0	4	44.4	19	137.53	96.42	6	31.6	1	5.3	0	0.0	1	14.3
Boys and girls	167	55	136.82	102.49	11	20.0	5	9.1	1	1.8	6	35.3	38	138.00	98.76	9	23.7	2	5.3	1	2.6	3	25.0
No. 14 Not Carious First Examination																							
Boys	86	27	138.00	101.07	5	18.5	1	3.7	0	0.0	1	16.7	16	135.19	103.38	3	18.8	0	0.0	0	0.0	0	0.0
Girls	81	22	135.36	104.95	4	18.2	1	4.5	0	0.0	1	20.0	21	134.24	100.95	5	23.8	0	0.0	0	0.0	0	0.0
Boys and girls	167	49	136.82	102.82	9	18.4	2	4.1	0	0.0	2	18.2	37	134.65	102.00	8	21.6	0	0.0	0	0.0	0	0.0
No. 19 Not Carious First Examination																							
Boys	86	20	142.25	96.85	4	20.0	0	0.0	1	5.0	1	20.0	18	138.50	98.22	8	44.4	1	5.6	0	0.0	1	11.1
Girls	81	17	132.35	103.47	4	23.5	0	0.0	2	11.8	2	33.3	10	134.30	98.00	2	20.0	1	10.0	0	0.0	1	33.3
Boys and girls	167	37	137.70	99.89	8	21.6	0	0.0	3	8.1	3	27.3	28	137.00	98.14	10	35.7	2	7.1	0	0.0	2	16.7
No. 30 Not Carious First Examination																							
Boys	86	21	139.05	101.14	6	28.6	1	4.8	0	0.0	1	14.3	20	137.55	103.00	7	35.0	0	0.0	1	5.0	1	12.5
Girls	81	13	136.92	101.23	2	15.4	1	7.7	1	7.7	2	50.0	13	134.62	103.85	4	30.8	3	23.1	0	0.0	3	42.9
Boys and girls	167	34	138.24	101.18	8	23.5	2	5.9	1	2.9	3	27.3	33	136.39	103.33	11	33.3	3	9.1	1	3.0	4	26.7

measured by the rather slow process of dental decay. Substantial equality of the two groups in this respect, therefore, is perhaps just what should be expected.

A more delicate measure of the effect of instruction is the amount of dental care which was given the teeth which needed such care. This was measured by recording the number of teeth which were in perfect condition at the beginning of the experiment and became carious during the experiment. These teeth evidently presented new problems in dental care. To determine how effectively these problems were met, a record was further kept of the number of such teeth which were filled

TABLE VII.—SUMMARY OF NUMBER OF FIRST PERMANENT MOLARS BECOMING CARIOUS DURING THE EXPERIMENTAL PERIOD WHICH WERE GIVEN DENTAL CARE

	Summary of first permanent molars			Second examination					
	Not carious first examination	Carious		Filled		Extracted		Received dental care	
		No.	Per cent	No.	Per cent	No.	Per cent	No.	Per cent
Experimental school.									
Boys	95	21	22.1	3	3.2	2	2.1	5	19.2
Girls	80	15	18.8	6	7.5	3	3.8	9	37.5
Boys and girls	175	36	20.6	9	5.1	5	2.9	14	28.0
Control school.									
Boys	73	21	28.8	2	2.7	2	2.7	4	16.0
Girls	63	17	27.0	5	7.9	0	0.0	5	22.7
Boys and girls	136	38	27.9	7	5.1	2	1.5	9	19.2

or extracted during the period of the experiment. This record is tabulated for each of the four six-year molars separately. This tabulation is presented in Table VI. The individuals are selected from the groups which were equalized by age and approximately by average IQ. It may be read as follows: Among the eighty-six boys who made up the equalized group in the experimental group there were twenty-seven perfect teeth on the first examination. On the second examination it was found that there were six carious teeth, one filled and one

extracted. In other words, eight teeth had become carious and of these two had received dental care. Twenty-five per cent of the teeth which had become carious, therefore, had received dental care. The significant figures, therefore, are the percentages under the column, Received Dental Care.

The main facts of Table VI are summarized in Table VII. It appears, for example, that among the boys of the experimental group there were ninety-five six-year molars which were not carious at the first examination. Of these, twenty-six became carious, and twenty-one of these twenty-six remained carious at the second examination. Three had been filled and two had been extracted. Of the teeth which had become carious, therefore, 19.2 per cent had received dental care. Among the boys in the control group, twenty-five teeth had become carious, and of these, four had received dental care, or 16 per cent. In the case of the girls of the experimental group, 37.5 per cent of the teeth which became carious received dental care and of the girls in the control group 22.7 per cent of the teeth which had become carious received dental care. The per cent for boys and girls together in the case of the experimental group is 28 and the control group 19.2. While the numbers are small, the evidence points in the direction of the conclusion that the children of the experimental group received better dental care than did those of the control group. In neither group, however, can the instruction be regarded as having been very effective, since fewer than 30 per cent of the teeth which required dental care received it.

Summary

The experiment was carried on in two fifth and two sixth grades of two public schools. Group 1 which saw the films was in one school and Group 2 which did not see the films was in the second school. The purpose was to measure the effect of seeing the motion pictures on the children's behavior.

The procedure consisted in giving the same instruction for thirteen school days to the parallel groups. The teaching was done by a trained assistant. The subject-matter included instruction on the structure and development of the teeth, on diet, and the care of the teeth. The same detailed outline was followed with both groups. With Group 1 oral instruction was supplemented by pictures, diagrams, and models, and also by two motion picture films entitled, Tommy Tucker's Tooth and Clara Cleans Her Teeth. With Group 2 the

motion picture film was not used. The number of children in Group 1 was 208 and in Group 2, 182.

The results of the instruction were measured by:

1. An information test.
2. A questionnaire in which the children reported on the care of their teeth at home and the food which they ate.
3. Material brought in by the children.
4. Condition of the teeth as determined by dental examination.

1. *Information Test.*—The groups were first equalized on the basis of intelligence test scores. Group 1 advanced from a score of 17.03 points to 29.47 points while Group 2 advanced from 14.85 points to 29.69 points. Group 1 thus advanced 12.44 points and Group 2, 14.84 points or 2.4 points more. The film group thus gained less on the information test than the control group.

2. *Questionnaire.*—The children of both groups reported almost universal use of the toothbrush and eating of vegetables, fruit, and milk. Fewer children of Group 2 reported using the toothbrush daily (eighty-one per cent as compared with ninety-eight per cent). The difference seems hardly significant.

3. *Material Brought In and Read.*—Group 2 read about the same number of booklets and took home more literature, but Group 1 brought in many more clippings and contributed more original articles, cartoons, and pictures. This comparison is somewhat in favor of the film group.

4. *Results of Dental Examination.*—Since age is a factor in the development and decay of the teeth, the groups were equalized by ages. After such equalization the comparison was made of the average number of carious six-year-old molars of the two groups on examinations made before and after instruction. In Group 1 the average number per child increased from 1.46 to 1.57 and in Group 2 the average number decreased from 1.90 to 1.80. The difference may have been due to a difference in the number of carious teeth on the first examination. The groups were then equalized on the basis of the condition of their six-year-old molars on the first examination. In Group 1 after equalization the average number per child of carious six-year-old molars decreased from 2.03 to 1.79, and in Group 2 the average number decreased from 2.02 to 1.87. This shows a slight advantage for the film group.

A comparison was then made of the molars which were in perfect condition on the first examination. A smaller number of perfect

molars became carious in the case of Group 1 than in the case of Group 2, (20.6 per cent against 27.9 per cent). A larger number of these teeth received dental care in Group 1 than in Group 2, (28.0 per cent against 19.2 per cent). This comparison is slightly in favor of Group 1.

The children who did not see the film apparently gained more information from the instruction than those that did. The children who saw the film, however, appear to have been stimulated to somewhat greater activity in bringing in supplementary material. After careful analysis has been made of the data, the group which saw the film appears to have improved somewhat more in the condition of their teeth than the pupils who did not see the film. The difference, however, is not conclusive. A slightly larger percentage of the group which saw the film received dental care, but fewer than 70 per cent of the defects of either group were corrected. Further experimentation would be desirable with somewhat more carefully controlled conditions and with films which may be more effective than those which were used in the experiment.

INTELLIGENCE, MOTIVATION, AND ACHIEVEMENT

AUSTIN H. TURNEY

University of Kansas

Leading authorities in the field of psychology and of educational psychology recognize the existence of fundamental motivating mechanisms. Thus Gates writes of certain "cravings" and "urges"[1] and Woodworth of "dependable motives."[2] These fundamental motives are factors in school achievement to a greater or lesser degree, depending in part upon the method of teaching being used. Thus we have many writers accepting the project method as psychologically sound because it enlists certain of these dependable motives.[3] The literature of method is full of discussions concerning "motivation" or "motivating the pupil."

The beginning student in educational psychology sometimes fails to appreciate the direct and practical connection between fundamental motives and actual schoolroom work. It has been our experience that the importance of motivating mechanisms is made more clear if the treatment of this topic is linked up with the discrepancy between intelligence and achievement. This approach not only emphasizes the importance of the rôle of motivation, but also recognizes the fact that one great educational problem is that of determining the extent to which motivating mechanisms affect achievement, and of determining the possibility of affecting these mechanisms through teaching technique and choice of content.

A number of investigators have submitted data indicating that the discrepancy between intelligence test results and achievement usually found in schoolroom situations is not necessarily the result of faulty measures either of intelligence or of achievement. On the contrary, this discrepancy is the natural result of other factors, chief among which are certain traits or types of behavior which for want of

---

[1] Gates, A. I.: "Psychology for Students of Education." The Macmillan Company, 1930, Chap. VI.

[2] Woodworth, R. S.: "Psychology." Henry Holt and Co., 1929, Chap. VI.

[3] Jordan, A. M.: "Educational Psychology." Henry Holt and Co., 1929, p. 113.

Kilpatrick, W. H.: "Foundations of Method." The Macmillan Company, 1926.

Douglass, H. R.: "Modern Methods of High School Teaching." Houghton Mifflin, 1926, Chap. XI.

better terms we may call "industry," "persistence," "ambition," "school attitude," and "dependability."[1]

This paper will present data suggesting that the behavior of pupils characterized by such terms as "industry," "perseverance," "dependability," and "ambition," as judged by teachers is probably an index of conditions of motivation, and indicating also the importance of the factors of motivation in school achievement. The data were secured in a study carried out in the University of Minnesota High School. Every student in each of the four classes in the High School was rated by each of his teachers twice during the school year on each of nine traits. The reliability of the judgments, as found by correlating first and second judgments made about sixty days apart, was, on the average, .80.

Intercorrelations were calculated for the following variables:[2]

1. Numerical value of all marks earned by each student during his entire residence in the high school
2. Chronological age
3. IQ derived from five or more tests
4. Mental age
5. Industry
6. Perseverance
7. Dependability
8. Ambition

For the Junior and Senior classes partial correlations of the first order were calculated holding constant in turn CA, MA, IQ, and marks.

Table I presents zero order correlations concerning the Junior class. Attention is especially called to the following groups of coefficients showing the correlations between certain variables:

Marks and MA	.57	MA and industry	.33
Marks and IQ	.61	MA and perseverance	.35
Marks and industry	.80	MA and dependability	.37
Marks and perseverance	.80	MA and ambition	.40
Marks and dependability	.82	Industry and perseverance	.96
Marks and ambition	.84	Industry and dependability	.96
IQ and industry	.37	Industry and ambition	.93
IQ and perseverance	.48	Perseverance and dependability	.95
IQ and dependability	.40	Perseverance and ambition	.91
IQ and ambition	.44	Dependability and ambition	.93

[1] See Turney, A. H.: "Factors Other than Intelligence That Affect Success in the High School as Indicated by Teachers' Marks." Minneapolis, Minn., University of Minnesota Press, 1930, pp. 31ff. for a review of these studies.

[2] The original data involve six other variables omitted as not pertinent to this analysis.

TABLE I.—INTERCORRELATIONS BETWEEN VARIABLES INDICATED
(N = 65 Juniors, University of Minnesota High School, 1926–1927)

	1 Marks	2 CA	3 IQ	4 MA	5 Industry	6 Perseverance	7 Dependability	8 Ambition
1. Marks	.....	−.314	.608	.569	.798	.798	.822	.842
2. CA	.....	.....	−.493	−.419	−.201	−.307	−.118	−.358
3. IQ	.....	.....	.....	.952	.365	.481	.397	.444
4. MA	.....	.....	.....	.....	.326	.352	.368	.398
5. Industry	.....	.....	.....	.....	.....	.964	.956	.928
6. Perseverance	.....	.....	.....	.....	.....	.....	.950	.913
7. Dependability	.....	.....	.....	.....	.....	.....	.....	.934
8. Ambition	.....							

Table II presents similar data concerning the Senior class in which attention is invited to the following correlations:

Marks and MA	.67	MA and industry	.35
Marks and IQ	.69	MA and perseverance	.31
Marks and industry	.66	MA and dependability	.41
Marks and perseverance	.63	MA and ambition	.39
Marks and dependability	.72	Industry and perseverance	.96
Marks and ambition	.68	Industry and dependability	.90
IQ and industry	.33	Industry and ambition	.78
IQ and perseverance	.32	Perseverance and dependability	.91
IQ and dependability	.41	Perseverance and ambition	.83
IQ and ambition	.40	Dependability and ambition	.82

TABLE II.—INTERCORRELATIONS BETWEEN VARIABLES INDICATED
(N = 48 Seniors, University of Minnesota High School, 1926–1927)

	1 Marks	2 CA	3 IQ	4 MA	5 Industry	6 Perseverance	7 Dependability	8 Ambition
1. Marks	.....	−.315	.691	.671	.661	.630	.717	.675
2. CA	.....	.....	−.736	−.712	−.255	−.202	−.265	−.390
3. IQ	.....	.....	.....	.986	.330	.315	.407	.399
4. MA	.....	.....	.....	.....	.348	.305	.405	.389
5. Industry	.....	.....	.....	.....	.....	.956	.895	.780
6. Perseverance	.....	.....	.....	.....	.....	.....	.911	.834
8. Dependability	.....	.....	.....	.....	.....	.....	.....	.815
7. Ambition	.....							

For our purposes we may summarize the two tables as showing a correlation between marks and IQ slightly higher than the average coefficient of correlation between measures of intelligence and measures of achievement reported in the literature.[1] Between marks and the traits of "industry," "perseverance," "dependability" and "ambition," the correlations are as high and, more often than not, higher than the correlation between IQ and marks or that between MA and marks The correlation between IQ and each of these four traits and between. MA and each of these four traits is relatively low, but the correlation between each pair of these four traits is high.

TABLE III.—INTERCORRELATIONS BETWEEN VARIABLES INDICATED WITH CA CONSTANT

(N = 65 Juniors, University of Minnesota High School, 1926–1927)

	1 Marks	2 IQ	3 MA	4 Industry	5 Perseverance	6 Dependability	7 Ambition
1. Marks	....	.546	.506	.789	.776	.831	.823
2. IQ	....	....	.942	.311	.398	.392	.330
3. MA	....	....	....	.272	.258	.354	.293
4. Industry	....	....	....	....	.969	.958	.936
5. Perseverance	....	....	....	....	....	.968	.905
6. Dependability	....	....	....	....	....	....	.963
7. Ambition							

The relatively low correlation between each of the traits and either IQ or MA suggests that neither are these "intellectual" traits nor are the judgments the effect of halo. Halo effect of marks upon judgments of the traits can scarcely be present since each of the judges furnished only one-twelfth of the Junior marks and only one-sixteenth of the Senior marks.

In Tables III and IV are the partial coefficients of correlation between the variables with CA constant. Of the correlations especially pointed out above the greatest change is in the case of that between IQ and marks which is somewhat reduced. In the main, the correlations between IQ and marks, IQ and the four traits, marks and the four traits, and between the traits themselves are not greatly affected when CA is held constant.

[1] Turney, A. H.: *Op. cit.*, Table I.

Tables V and VI present the partial coefficients of correlations between each of the variables with IQ constant. It is to be noted that the correlations between marks and each of the four traits, "industry,"

TABLE IV.—INTERCORRELATIONS BETWEEN VARIABLES INDICATED WITH CA CONSTANT
(N = 48 Seniors, University of Minnesota High School, 1926–1927)

	1 Marks	2 IQ	3 MA	4 Industry	5 Perseverance	6 Dependability	7 Ambition
1. Marks	....	.721	.670	.635	.610	.695	.633
2. IQ	....	....	.977	.219	.252	.327	.181
3. MA	....	....	....	.635	.610	.695	.633
4. Industry	....	....	....	....	.956	.889	.766
5. Perseverance	....	....	....	....	....	.909	.837
6. Dependability	....	....	....	....	....	....	.803
7. Ambition							

"perseverance," "dependability" and "ambition," and also those between the traits themselves are but slightly changed.

TABLE V.—INTERCORRELATIONS BETWEEN VARIABLES INDICATED WITH IQ CONSTANT
(N = 65 Juniors, University of Minnesota High School, 1926–1927)

	1 Marks	2 CA	3 MA	4 Industry	5 Perseverance	6 Dependability	7 Ambition
1. Marks	.....	−.020	−.283	.783	.728	.800	.803
2. CA	.....	.....	.184	.026	−.092	.098	−.176
3. MA	.....	.....	.....	−.072	−.387	−.035	−.089
4. Industry	.....	.....	.....	.....	.964	.950	.937
5. Perseverance	.....	.....	.....	.....	.....	.944	.888
6. Dependability	.....	.....	.....	.....	.....	.....	.921
7. Ambition							

The results of holding constant MA are shown in Tables VII and VIII. The correlations between marks and the traits, and between the four traits themselves are but slightly changed.

When marks are held constant, as shown in Tables IX and X, the correlations between the four traits, "industry," "perseverance,"

TABLE VI.—INTERCORRELATIONS BETWEEN VARIABLES INDICATED WITH IQ CONSTANT
(N = 48 Seniors, University of Minnesota High School, 1926–1927)

	1 Marks	2 CA	3 MA	4 Industry	5 Perseverance	6 Dependability	7 Ambition
1. Marks	.....	.398	−.098	.634	.601	.661	.602
2. CA	.....	....	.147	−.019	.047	.057	−.156
3. MA	.....	....	.....	.173	−.045	.031	−.031
4. Industry	.....	....	.....	.....	.953	.884	.749
5. Perseverance	.....	....	.....	.....	.....	.906	.816
6. Dependability	.....	....	.....	.....	.....	....	.757
7. Ambition							

"dependability" and "ambition," and both MA and IQ become very low and usually negative. The correlations between the traits themselves are still high.

TABLE VII.—INTERCORRELATIONS BETWEEN VARIABLES INDICATED WITH MA CONSTANT
(N = 65 Juniors, University of Minnesota High School, 1926–1927)

	1 Marks	2 CA	3 IQ	4 Industry	5 Perseverance	6 Dependability	7 Ambition
1. Marks	.....	−.097	.257	.790	.777	.803	.818
2. CA	.....	.....	−.333	−.075	−.188	.043	−.229
3. IQ	.....	.....	.....	.187	.500	.203	.227
4. Industry	.....	.....	.....	.....	.956	.953	.922
5. Perseverance	.....	.....	.....	.....	.....	.943	.901
6. Dependability	.....	.....	.....	.....	.....	....	.925
7. Ambition							

The foregoing data indicate that the four traits, "industry," "perseverance," "dependability" and "ambition" are more closely related to

achievement than is either IQ or MA, and that they are not closely related to either IQ or MA, but that they are so closely related to each

TABLE VIII.—INTERCORRELATIONS BETWEEN VARIABLES INDICATED WITH MA CONSTANT
(N = 48 Seniors, University of Minnesota High School, 1926–1927)

	1 Marks	2 CA	3 IQ	4 Industry	5 Perseverance	6 Dependability	7 Ambition
1. Marks	.....	.523	.286	.614	.602	.657	.605
2. CA	.....	....	−.343	−.011	.022	.036	−.174
3. IQ	.....	.....	.....	.098	.104	.062	.115
4. Industry	.....	....	.....	.....	.964	.883	.751
5. Perseverance	.....	....	.....	.....	....	.908	.821
6. Dependability	.....	....	.....	.....	....	....	.786
7. Ambition							

other as to show marked overlapping.[1] In view of the high reliability found for all of these data we can only conclude that the teachers were in fact actually judging accurately the behavior of their students in a

TABLE IX.—INTERCORRELATIONS BETWEEN VARIABLES INDICATED WITH MARKS CONSTANT
(N = 65 Juniors, University of Minnesota High School, 1926–1927)

	1 CA	2 IQ	3 MA	4 Industry	5 Perseverance	6 Dependability	7 Ambition
1. CA	.....	−.401	−.307	.088	−.098	.257	−.182
2. IQ	.....	.....	.931	−.252	−.008	−.227	−.158
3. MA	.....	.....	.....	.260	−.207	−.213	−.181
4. Industry	.....	.....	.....	.....	.908	.877	.784
5. Perseverance	.....	.....	.....	.....	.....	.860	.739
6. Dependability	.....	.....	.....	.....	.....	.....	.778
7. Ambition							

specific schoolroom situation where these traits represent the extent to which the students were motivated to conform to classroom require-

[1] Garrett, H. E.: "Statistics in Psychology and Education." Longmans, Green & Co. 1926, p. 291.

ments and to achieve in classroom activities. Hence these judgments represent largely the effects of motivation and they necessarily overlap as already shown. This overlapping of the traits and the lack of correlation between these traits and either IQ, MA, or CA point to the conclusion that the motivation existing in this situation was in reality the effect of "drive" or "dependable motives" as affected by the situation.

The significance and implication of these data are, we believe, rather great. They should serve to focus attention upon the real nature of the discrepancy between "intelligence" and school achieve-

TABLE X.—INTERCORRELATIONS BETWEEN VARIABLES INDICATED WITH MARKS CONSTANT

(N = 48 Seniors, University of Minnesota High School, 1926–1927)

	1 CA	2 IQ	3 MA	4 Industry	5 Perseverance	6 Dependability	7 Ambition
1. CA	.....	−.026	−.712	.066	−.007	−.059	−.255
2. IQ	.....	.....	.972	.223	−.211	−.175	−.126
3. MA	.....	.....	.....	−.172	−.204	−.148	−.118
4. Industry	.....	.....	.....	.....	.943	.808	.606
5. Perseverance	.....	.....	.....	.....	.....	.851	.719
6. Dependability	.....	.....	.....	.....	.....	.....	.650
7. Ambition							

ment. It seems clear that the two major factors in school achievement are intelligence on the one hand and "motivation" on the other. But motivation is not to be regarded as the product entirely of the immediate situation. The presence of fundamental driving mechanisms, the play of interests developed long before, and the varying effect of the immediate situation in the schoolroom are probably all a part of what we call "motivation."

Our data serve only to emphasize the presence and importance of factors of motivation. It would be desirable to know to what extent these factors are the result of specific environmental situations such as the classroom. To what extent can the "dependable" motives be influenced by specific schoolroom situations such as the use of the project method or the socialized recitation? In other words to what extent can the industry and ambition of pupils be changed through educational influences? The variability in responsiveness to specific

stimulating situations in the classroom may be as much a biologically fixed quality as "intelligence" is usually considered to be.[1] On the other hand, it may be vastly more subject to educational (*i.e.*, environmental) influences.

If we accept for purposes of discussion the concept of *g* as set forth by Spearman[2] the situation can be presented in the following way: Two students possessing equal amounts of *g* would not necessarily utilize it equally in the same classroom activity. Moreover, even objectively measured, achievement must involve a great deal more than *g*. Not all classroom activities would correlate highly with measures of *g*.[3] Many of them would depend as much or more upon just these characteristics of the student which we label "industry," or "perseverance" for success. Hence the correlations shown between "industry," "perseverance," "dependability" and "ambition," and marks, are probably measures of a true relationship. To the extent that the classroom work involves activities not largely dependent upon a pupil's *g*, or to the extent that activities very largely involving *g* are not attended to in equal proportion by students approximately equally equipped, a teacher's marks could not correlate perfectly with measures of intelligence.

These facts might well be borne in mind in using mental tests. They do not minimize the utility of tests of "intelligence" but rather enhance their value. In ability grouping, for example, until the possibilities of motivation are known, we must agree with Billett,[4] who is one of the few to point out the advisability of grouping on measures of mental ability alone. To include in the criteria for grouping measures of past achievement covers up the very important effects of motivation. The possibilities of educational guidance, and the understanding of difficulties are much greater with the use of the single criterion. If criteria such as achievement tests which measure, in part, the effects of motivation are used, they should be utilized in a manner permitting the teacher's knowledge of their presence and weight.

---

[1] I am not unaware of the arguments against this concept, as in the Twenty-seventh Yearbook of the National Society for the Study of Education, Part I, but see Jennings, H. S.: "The Biological Basis of Human Nature." New York: W. W. Norton, 1930, p. 24.

[2] Spearman, C.: "The Abilities of Man." The Macmillan Co., 1927.

[3] Ibid., Chap. XII.

[4] Billett, R. O.: A Controlled Experiment to Determine the Advantages of Homogeneous Grouping. Ohio State University, *Educational Research Bulletin*, Vol. VII; Nos. 7, 8 and 9, April 4, 18, May 2, 1928.

THE "OMISSION" AS A SPECIFIC DETERMINER IN THE TRUE-FALSE EXAMINATION

C. C. WEIDEMANN

Teachers College, University of Nebraska

Do students tend to omit items of the true-false examination that are *true* more often than they tend to omit items that are *false?* In this study each item is either true or false by published authority as the criterion.

The *specific determiner*[1] is a word or phase which occurs in the *external form* of a test item and alters the *internal constitutiom* of that test item so that the correct response is either more often *true* than *false* or more often *false* than *true.* The *specific determiner* in any given test item may be so controlled by the individual who constructs the true-false examination that the correct response is approximately either as often *true* as *false* or as often *false* as *true.*

Wood[2] differentiates between *internal constitution* and the *external form* of examination items. He says:

> The essential character of a question, and consequently its "statistical behavior" depends upon its content as well as its external form. The content of a question is often very subtle; many of the "simplest" questions are found on inspection to be compounds of many subtleties. It seems quite reasonable to suppose that within the limits of accepted forms of questions, the "inner" nature of a question is more important than its "outer" form; although as a matter of common sense, one may say that the "outer" aspect exerts more influence as we go down the intellectual or educational ladder and less as we ascend it.

Assume that the *specific determiners* of a true-false examination have been controlled, yet the variable of the temperament of the examinees is still present. The temperament of examinees may cause some to respond more often *true* than *false* and other to respond more often *false* than *true* to the same test item in a given true-false examination. This variable of temperament of the examinee seems to be somewhat a specific determiner in relation to any given test item. A study of this point needs to be made.

---

[1] Weidemann, C. C.: "How to Construct the True-false Examination." Bureau of Publications, Teachers College, Columbia University, 1926, p. 68.

[2] Wood, Ben D.: Studies in Achievement Tests. *Journal of Educational Psychology,* Vol. XVII, January, February, 1926, pp. 1–22, 125–139.

So far as an individual examinee is concerned, no direct technique seems to be available to control the temperament of that examinee while responding to the items in true-false examinations whose *specific determiners* so far as *external form* are concerned, have been controlled on the assumption that *score equals right minus wrong.* In the case of large groups of examinees temperament might be considered to be a chance factor, compensating in such a way that the tendency for some examinees to respond more often *true* than *false* would be offset by the tendency of others to respond more often *false* than *true.* This study assumes the factor of temperament of the examinees to be a chance factor.

Fritz[1] and a few others have claimed that students tend to respond to true-false items more times as *true* than as *false.* The figures are sixty-two per cent *true* and thirty-eight per cent *false,* with technical information quite unknown to the students. He reports practically the same result with information which was studied by the students as a regular part of the class work.

Ruch[2] reports "a number of studies similar to that of Fritz" and gives fifty-two per cent marked as *true* and forty-eight per cent marked as *false.* In an unpublished study by D. D. Durrell, and C. L. Cushman, (data available through G. M. Ruch, University of California) "the results indicate that pure guessing, when it occurs, is roughly a 50:50 situation."

Rutledge[3] found "that when the 'same' items were stated in both true and false forms, they were equally difficult."

These three studies make no mention of the "omission" as a specific determiner.

In the fall of 1926 in a course entitled "Foundations of Modern Education," the writer used seven true-false examination tests based upon F. P. Graves's, "A Student's History of Education." At the beginning of the semester, the class totalled three hundred thirty-eight divided into three sections of ninty-nine, one hundred three, and one hundred thirty-six students respectively. A total of three hundred

---

[1] Fritz, M. F.: Guessing in a True-false Test. *Journal of Educational Psychology,* Vol. XVIII, 1927, pp. 558–561.

[2] Ruch, G. M.: "The Objective or New-type Examination." Scott, Foresman and Company, New York, 1929, pp. 365.

[3] Rutledge, R. E.: "The True-false Examination in Elementary Psychology, with Suggestions for Its Improvement." Unpublished Ph.D. Thesis, University of California, 1927.

thirty-one students included in this study remained throughout the semester. The members in each section were given a reading assignment from Graves and within two weeks were given a true-false test based upon the reading assignment. The object was primarily to measure immediate retention rather than delayed retention. No test contained any items which aimed to review material covered in any previously given test.

In order to reduce cheating to a minimum each of the seven bi-weekly tests was administered in either Forms A, B and C or Forms A and B. The students seated in Columns 1, 2, 3 had Forms A, B and C respectively; Columns 4, 5 and 6 had Forms A, B and C respectively, and so on. After the forms were answered they were collected by the instructor. Then Columns 1, 2 and 3 had Forms B, C, and A respectively; Columns 4, 5 and 6 had Forms B, C and A respectively, and so on. After the second order of Forms were answered they were collected by the instructor. Then Columns 1, 2, and 3 had Forms C, A and B respectively; and Columns 4, 5, 6 had Forms C, A, and B respectively, and so on. A similar procedure was followed when the bi-weekly test consisted of Forms A and B. The total test time of each bi-weekly test was from twenty-five to forty minutes.

The frequency with which each item of each Form of a given bi-weekly test was omitted was tabulated for each student taking each test. All results are reported upon the basis that the number of *true* items equals the number of *false* items in each of the seven bi-weekly tests. The *specific determiners* of "all and never," "more than, less than," etc., and "If . . . then" were controlled on the assumption that the *score equals rights minus wrongs.*

The directions commanded the student to mark with a plus (+) only those statements which were entirely *true;* if the statement were partly or entirely *false* it was marked with a zero (0). A response space was provided to the left of each statement. The students were told *not* to guess.

Table I displays the number of times each omitted item was either *true* or *false.* A grand total of 3064 students responded to a grand total of 495 true-false statements, accumulating a grand total of 3954 omissions, in seven bi-weekly tests administered during one semester. Table I shows that forty-four per cent of the omitted statements were *true* and fifty-six per cent of the omitted statements were *false.* Is the twelve per cent difference significant? The critical ratio value is sixteen, so that the difference is significant.

TABLE I.—NUMBER OF TIMES *Omitted* ITEMS ARE EITHER *True* OR *False* FOR EACH OF THE SEVEN TRUE-FALSE TESTS

Bi-weekly true-false test number	The number of times test items were omitted by the student, which items were		Total omissions
	True	False	
1	2	3	4
1	737	655	1392
2	270	283	553
3	159	400	559
4	219	203	422
5	57	46	103
6	112	229	341
7	172	412	584
Totals.....	1726	2228	3954
Per cent.....	44	56	100

Even though the statistical per cent difference is significant, there is relatively little value in the per cent difference so far as affecting the score and placement of an individual student is concerned.

Assume:

1. The score equal to the total number of items minus the *omissions* minus two times the wrongs, *i.e.*, $S = T - O - 2W$.
2. That on the average not more than twenty items out of every one-hundred items of a true-false examination would be omitted by a given student. Some may omit more and others less items.
3. That every item so omitted would then be marked *false.*

Then upon the basis of this study, fifty-six out of every one hundred omitted items arbitrarily marked false would be *right* and forty-four would be *wrong.* $S = 56 - 44 = 12$. But, it is assumed that on the average, not more than twenty items out of every one hundred items of a true-false examination would be omitted and arbitrarily marked *false.* Thus the contribution of the *omission* 12×20 per cent or 2.4 points of score. This change of score seems administratively to be relatively insignificant in the degree to which it may affect the placement of an individual in a group.

1. Three hundred thirty-one students who took the seven bi-weekly true-false tests based upon Graves', "A Student's History of Educa-

tion," in crude per cents omitted items that were *false* more often that they omitted items that were *true*.

2. The *omission* in the seven bi-weekly true-false examinations of this study is a statistically significant *specific determiner*. The critical ratio value is 16.

3. It seems to the writer that in a very carefully constructed true-false examination "omissions" are probably as often *true* as *false*, and as often *false* as *true*.

4. From the standpoint of affecting the grade placement of an individual student in a group, this study would tend to indicate that the *omission* as a specific determiner is not significant.

5. Further study will probably show the *omission* of the true-false examination to be statistically insignificant.

FACTORS INVOLVED IN CHILDREN'S FRIENDSHIPS[1]

GLADYS GARDNER JENKINS

University of Chicago

The following study was made in the effort to discover definite factors which influence the forming of friendships among children. There are, of course, many factors of a personal nature not yet possible of analysis, but there are others such as the relative intelligence of child and friend, age differences, play interests, and like factors which are measurable in some degree and may bear upon the question. The study included two hundred eighty boys and girls representing a cross section of the junior high schools of Riverside, California, a city of approximately 35,000 inhabitants.

Chronological Age

The chronological age of one hundred eighty children was compared with that of their friends. The range for one hundred twenty-five girls was twelve years and five months to sixteen years and eleven months; for their friends twelve years to nineteen years and eight months. For fifty-five boys the age range was twelve years and three months to sixteen years and four months; for their friends twelve years and five months to seventeen years. The mean chronological age for the children was fourteen years and three months and for their friends fourteen years and five months, with a standard deviation of 10.6 months for the children, and 13.6 months for their friends.

Of the one hundred eighty boys and girls studied ninety-seven made their friendships in school, eighty-three outside of school. The mean chronological age for those friendships made in school was fourteen years and four months for both the children and their friends, with a standard deviation of 10.3 months for the children and eleven months for their friends. The mean chronological age for the eighty-three friendships made outside of school was fourteen years and two months for the children and fourteen years and six months for their friends, with a standard deviation of 10.9 months for the children and 16.1 months for their friends. As there is a larger spread in the chronologi-

[1] The writer wishes to express her indebtedness to the guidance of Dr. Frank N. Freeman in this study, and to former Superintendent A. N. Wheelock and Principal F. P. Taylor of the Riverside School System for their cooperation in making available their records.

cal age range for those friends which are made outside of school than for those made within the school, it would seem as if the school has a tendency to diminish the age difference in children's friendships. In spite of the larger standard deviation for the friendships made outside of school there is no uniform or perceptible tendency for children to choose friends either older or younger than themselves. The data are presented in Table I.

TABLE I.—COMPARISON OF CHRONOLOGICAL AGE DIFFERENCE BETWEEN CHILD AND FRIEND WITH THE PLACE WHERE THEY MET

Age difference, months	Neighborhood and places other than school		School	
		Per cent		Per cent
Within one month	6	9.6	5	4.8
1 to 6	22	31.9	40	38.8
7 to 12	17	24.7	46	44.7
13 to 18	12	17.5	6	5.9
19 to 24	4	5.8	3	2.9
25 to 30	4	5.8	3	2.9
31 to 36	1	1.4		
37 to 42	2	2.9		
43 to 48				
49 to 54				
55 to 60	1	1.4		

The Pearson product-moment correlation coefficient was $+.470 \pm .039$. The correlation coefficient for the ninety-seven friendships made within the school was $+.529 \pm .049$, for the eighty-three friendships made outside of school, $+.462 \pm .073$. It would therefore seem that the tendency for children to choose friends of their own age exists independently of school grouping although it may be increased by it.

INTELLIGENCE QUOTIENTS

Intelligence quotients obtained by the Terman Group tests within the two previous years were available from the school records for one hundred ninety-seven of the children, one hundred thirty-four girls and sixty-three boys. The correlation coefficient was $+.332 \pm .043$ as against a correlation coefficient of $+.470 \pm .039$ for the chronological age. The correlation coefficient for the intelligence quotients of one hundred eight friendships made within the school was $+.299 \pm .059$,

for eighty-nine friendships made outside of the school the correlation coefficient was $+.374 \pm .062$. Inasmuch as the correlation coefficient for the friendships made out of school is slightly higher than that for those made in school, even allowing for the equalization which might occur because of the probable error, the indications would seem to be that intelligence quotient is an independent factor in children's friendships, not simply the outcome of school sectioning.

The mean intelligence quotient for both the children and their friends was 109.8 with a standard deviation for both groups of 15.3. For those friendships made in school the mean intelligence quotient is 109.5 for the children, 111.1 for their friends, with a standard deviation of 14.1 for the children and 14.8 for their friends. For those friendships made outside of school the mean intelligence quotient is 110.1 for the children, 108.2 for their friends, with a standard deviation of 16.7 for the children and 15.7 for their friends.

Mental Age

The mental age of one hundred seventy-two children and their friends, of which one hundred twenty were girls, fifty-two boys, was computed for the time of the study rather than for the time at which the friendship was made, as it did not seem possible to secure an accurate record of the exact time when each friendship was formed. Inasmuch as all the children studied had been tested by the Terman Group test within the two previous years it was felt for the purpose of this study it would be satisfactory to assume that the intelligence quotients had remained constant and to compute the mental age from the intelligence quotient and the chronological age at the time of the study.

The mental age range for both the children and their friends was eleven years and one month to twenty-one years eleven months. Seven children had the same mental age as their friends, thirty-seven children chose friends with a mental age within six months of their own, sixty-two had a mental age more than six months higher than their friends, sixty-six more than six months lower. There seems to be an equal tendency for a child to choose a friend with a higher mental age as there is for him to choose a friend with a lower one. The mean mental age for the children was fifteen years and seven months with a standard deviation of 22.7 months, the mean mental age for their friends was approximately the same, fifteen years and ten months with

a standard deviation of 21.9 months. The correlation coefficient was $+.423 \pm .042$.

A comparison of the correlation coefficients for chronological age, mental age, and intelligence quotient reveals the following data in order of the probable importance.

TABLE II.—A COMPARISON OF THE CORRELATION COEFFICIENTS OF CA, MA AND IQ

Subject	Correlation coefficient	PE
Chronological age	+.470	±.039
Mental age	+.423	±.042
Intelligence quotient	+.332	±.043

SCHOOL DIVISIONS

The fact that children seem to tend toward a choice of friends of approximately the same chronological age and mental age would suggest that these friends might also tend to be in the same grade or section in school. The following facts bear out this supposition. Of two hundred fifty-four children studied (one hundred twelve boys, one hundred fifty-one girls) one hundred thirteen of their friends were in the same section, sixty-two in the same grade but not the same section, seventy-nine in other grades. Of the one hundred twelve boys twenty-three or twenty per cent chose friends in a higher grade, twenty-one or eighteen per cent in a lower grade, and seventy or sixty-one per cent in the same grade. Of the one hundred fifty-one girls eight or five per cent chose friends in a lower grade, thirty-two or twenty-one per cent in a higher grade, one hundred eleven or seventy-four per cent in the same grade. The boys seem to have a higher range of friendships than the girls, sixty-one per cent of the friends of the boys were in the same grade as against seventy-four per cent of the girls. This tendency, if it holds true with a larger group, may be due to the fact that there is a slight tendency for boys to make more friendships in the neighborhood and through clubs than girls and, as has been shown, the age range for friendships made outside of school tends to be slightly higher than for school-friendships.

PLACES IN WHICH THE FRIENDSHIPS DEVELOPED

In order to discover whether the school, the neighborhood, the club, or other environmental factors have the most effect upon the

making of friendships the following factors concerning the places where the friendships were made were investigated:

1. School.
2. Neighborhood.
3. Club—Girl Scouts, Boy Scouts, Y. M. C. A., Girl Reserves.
4. Church.
5. Miscellaneous—family contacts, special interests—music, etc.

Information was obtained from two hundred fifty-five children (one hundred sixty girls, ninety-five boys). Table III indicates the number of friendships made in each of these places.

TABLE III.—NUMBER OF FRIENDSHIPS IN RELATION TO THE PLACE WHERE THE CHILDREN MET

Place	Number of friends	Per cent
School	138	54
Neighborhood	63	25
Church	20	8
Club	8	3
Miscellaneous	26	10

The table shows that the school was the place in which the largest number of friendships was made, with the neighborhood ranking second. A slight tendency was found for the boys to make more of their friendships in the neighborhood and places other than school than for the girls to do so. Forty-nine per cent of the boys and forty-four per cent of the girls made their friendships in places other than school.

Table IV indicates that the section grouping has a distinct influence upon the friendships which children make within the school, whereas for those friendships made out of the school the section has very little significance.

TABLE IV.—RELATION OF SECTION GROUPING TO FRIENDSHIPS MADE IN THE SCHOOL AND OUT OF THE SCHOOL

School division of friend	Friendship made in school, per cent	Friendship made out of school, per cent
Same section	64	19
Same grade (not same section)	25	24
Other grade	11	57

Occupation of Parent

In order to attempt to determine the influence of the social-economic status of families upon the friendships of the children a comparison was made between the occupation of the father of each child with that of the father of his friend. For the purpose of the study the occupations were grouped into the following four groups, coinciding approximately with the social groupings of the community:

1. Unskilled and semi-skilled labor.
2. Skilled labor (carpenter, electrician, etc.).
3. Semi-professional (nurse, school teacher, etc.), commercial, owner of small ranch or citrus grove, clerical.
4. Professional (engineer, physician, etc.), business executive, owner of large ranch or citrus grove.

Two hundred thirty-seven cases were studied (ninety-five boys, one hundred forty-two girls). It must be remembered that California is a large middle class state, and therefore the total distribution of children is largely weighted toward the skilled labor and commercial groups. Table V indicates the results secured.

Table V.—Comparison of Occupation of Parent of Child with Parent of Friend

Parent's occupation		Unskilled		Skilled, etc.		Commercial, etc.		Professional, etc.	
Total number of cases	Per cent of total number	Number	Per cent	Number	Per cent	Number	Per cent	Number	Per cent
Unskilled labor									
16	7	11	69	4	25	1	6		
Skilled labor, etc.									
121	51	4	3	83	69	29	24	5	4
Commercial, etc.									
71	29	...	...	11	16	50	71	9	13
Professional, etc.									
30	13	...	...	1	3	7	23	22	74

These figures would seem to indicate that the social-economic status of the parents in the community is an important criterion in the choice of friends which the child makes. In each case the number of friends chosen from the same social-economic group equals more than sixty per cent of the total distribution:

	Per Cent
Unskilled labor	69
Skilled labor, etc	69
Commercial, etc	71
Professional, etc	74

The correlation coefficient $+.716 \pm .032$ strongly expresses this tendency for children to choose friends of the same social-economic group. As there were only four groups the correlation coefficient was corrected for coarse grouping by Shepherd's formula with a resulting correlation coefficient of $+.817$. The correlation coefficients for the boys and their friends and for the girls and their friends were practically the same, $+.695 \pm .036$ for the boy group, and $+.724 \pm .027$ for the girl group.

It seems probable that the significance of the high correlation coefficient is not greatly influenced by the proximity of the homes of the children and their friends, as only twenty-five per cent of the children stated that their friendships had been formed in the neighborhood. It might also be mentioned that the clubs and churches in Riverside are centralized rather than located in special residential districts, so that it is probable that those friendships made through these agencies were not unduly influenced by neighborhood divisions. There is also a suggestion of a tendency for children to reach into a higher social-economic group in the choice of a friend, forty-eight children chose friends in a higher group against twenty-three who chose them in a lower one. The facts given show, however, that the stronger tendency in a child's choice of a friend is toward those of the same social-economic level.

Play Interests

Lehman's Play Quiz was given to one hundred twenty-six children (twenty-eight boys, ninety-eight girls) and their friends. Unfortunately the test was given one week after the other data had been collected, and in the meantime a rearrangement of school groups had made it impossible to reach all the boys who had been studied at the original time. The number of boys, therefore, is too small to permit any definite statement concerning the findings but the trends will be of interest. In scoring the tests the number of activities carried on in the past week because the child liked to do them were checked against those of his friend, and the number of identical inter-

ests recorded. The child's interests were then checked against those of cases picked at random from the children studied, three cases for each of the girls, six for each of the boys. The mean number of like interests for child and friend, and child and other than friend is shown in Table VI. It would seem as if there were a slight tendency for the children and their friends to have the larger number of common interests.

TABLE VI.—MEAN NUMBER OF LIKE INTERESTS FOR CHILD AND FRIEND, AND CHILD AND OTHER THAN FRIEND

CHILDREN	MEAN NUMBER OF LIKE INTERESTS
Boys.	
Boy and Friend	18.5
Boy and other than friend	16
Girls.	
Girl and Friend	16.5
Girl and other than friend	14.5
Boys and Girls.	
Child and friend	17
Child and other than friend	15

A study was next made of Section D of the test in order to discover whether the child and best friend tend to have primary interests in common. Table VII shows that more children had a common interest in one of the three things they like to do best with their friend than with the children picked at random from the group. Twenty-seven per cent of the children studied had one of their three greatest interests in common with their best friend as compared with nineteen per cent who had one of their greatest interests in common with children selected at random.

TABLE VII.—A COMPARISON OF COMMON PRIMARY INTERESTS BETWEEN CHILD AND FRIEND, AND CHILD AND OTHER THAN FRIEND

Number of primary interests	Per cent of common primary interests between child and friend	Per cent of common interests between child and other than friend
1	27	19
2	5	1
3	1	

Summary of Conclusions

Social-economic Status of Parents.—Of the factors considered in this study primary importance is assigned to the social-economic status of the parents. The correlation coefficient for the social-economic position of parent of child and parent of friend was $+.716 \pm .032$. Corrected for coarse grouping by Shepherd's formula the coefficient was $+.817$. This high correlation does not seem to have been significantly influenced by the proximity of homes as only twenty-five per cent of the total number of children stated that they made their friends in the neighborhood. There was also a suggestion of a tendency for the children to choose friends of a higher social-economic group; forty-eight had friends in a higher group, twenty-three in a lower.

Chronological Age.—The correlation coefficient for chronological age was $+.470 \pm .039$. This is probably somewhat greater in those friendships made in school. Children tend to choose friends within one year of their age. There is no tendency to choose friends either older or younger. Friendships made in the neighborhood have a larger age range than friendships made in school.

Mental Age and Intelligence Quotient.—The correlation coefficient for mental age was $+.423 \pm .042$. The correlation coefficient for intelligence quotient was $+.332 \pm .043$. The fact that the intelligence quotient correlation coefficient does not seem to be influenced by school groupings would suggest a spontaneous tendency to choose friends of the same approximate intelligence.

School Divisions and Place of Meeting.—There is an almost equal division in the number of friendships made in the school and those made in the neighborhood or through home contacts, churches, and clubs. It must be recognized, however, that the school is the greatest single source with the neighborhood ranking second with twenty-five per cent of the total number of friendships. Of those friendships made in school the grade-section divisions seem to be of importance.

Play Interests.—According to the data secured from Lehman's Play Quiz there seems to be a slight tendency for children to have a greater number of like interests with their best friends than with other children. It can not be stated as to whether this common interest was the cause or outcome of the friendship.

A STUDY OF CLASSROOM BEHAVIOR

WILLARD C. OLSON

University of Michigan

Considerable interest has been shown in recent years in the development of techniques for the quantitative expression of personality traits not easily amenable to measurement by test methods. By the use of carefully defined categories of overt behavior, relatively short time samples, repeated observations, and systematic recording, the writer has attempted to accomplish the task of measurement through direct observation.[1] The adaptability of the general method to a variety of behavior traits has been illustrated by Goodenough.[2] The present investigation was initiated to test further the applicability of the method and to acquaint a class in the psychology of personality with its use.

Whispering was chosen as the behavior to be observed. Wickman[3] has shown that whispering ranks first in frequency of occurrence among behavior problems in children, although teachers do not regard its occurrence in a child as a serious behavior disorder. A behavior problem is defined frequently as a discrepancy between the environmental demands made on an individual and his adjustment to them. Whispering, according to this definition, may or may not be a behavior problem in a particular classroom depending upon the requirements made with respect to it. The present report is concerned with an attempt to measure and describe whispering behavior in schoolroom situations quite apart from any judgment concerning its desirability or undesirability.

Collection of Data

Arrangements were made to permit the class to spend one hour, from 9:00 to 10:00 A. M., in observation in an elementary school.[4]

---

[1] Olson, Willard C.: "The Measurement of Nervous Habits in Normal Children." Minneapolis: The University of Minnesota Press, 1929.

[2] Goodenough, Florence L.: Measuring Behavior Traits by Means of Repeated Short Samples. *Journal of Juvenile Research*, Vol. XII, 1928, pp. 230–235.

[3] Wickman, E. K.: "Children's Behavior and Teachers' Attitudes." New York: The Commonwealth Fund Division of Publications, 1928.

[4] The writer is indebted to Miss Katherine G. Young, Principal of the Marcy School, Minneapolis, for cooperation in this project.

Each observer was supplied with a mimeographed sheet of instructions giving explicit directions for making the observations.[1] A whisper for the purpose of the study was defined as an unauthorized oral communication—whispered or aloud—to a person or persons other than the teacher. The unit of measurement was defined as one or more acts of whispering per five-minute period. Two observers were stationed in each room at opposite sides and all rooms in the building studied at the same hour. One student substituted a series of observations made on a twelfth grade class in economics. Ten successive observations were made and recorded in a prescribed manner on seating charts of the rooms. The total number of time samples in which a child was recorded as whispering was called his *whisper score.*

Reliability of Whisper Scores

The reliability of the total score for ten five-minute records was obtained by correlating the values obtained on each child by the two observers in each room. The coefficient of correlation between the two sets of scores varied between .20 and .78. The predicted reliabilities (Spearman-Brown formula) for twenty five-minute periods varied between .33 and .88 for the grade groups studied (Table I). The reliability of the single observer of a high school class was studied by correlating the sum of the odd numbered observations with the sum of the even numbered observations. The reliability coefficient for five observations was found to be .67 with .80 the predicted value for the ten observations.

In the room where the lowest reliability coefficient was found, the method of teaching called for frequent shifts in the seating arrangement. Since the observers did not know the children except as a seat location, such changes introduced difficulties. A greater amount of time would have to be spent in such a situation, the room program slightly modified, or the children identified as individuals, to secure the requisite reliability. It is probable that some initial training in the use of the method would improve the accuracy of the observers. Increased reliability would be obtained by making a larger number of observations.

It was felt that the records obtained showed sufficient consistency to warrant further inquiry into their distribution and significance. In the subsequent analysis, the average of the records of two observers

[1] Copies of the instruction sheet may be obtained from the writer upon request.

was taken as the score for each child in the elementary school. The high school records are the results for a single observer.

TABLE I.—RELIABILITY OF WHISPER SCORES

Group	*N*	*r*, 10 obs.	*r* (Sp. Br.), 20 obs.
1*B*	32	.20	.33
3*B*–2*A*	35	.78	.88
3*A*–3*B*	34	.76	.86
4*A*–4*B*	40	.60	.75
5*B*–4*A*	35	.43	.60
6*B*	31	.29	.45
6*A*	30	.35	.52
12	27	.80	.89

DISTRIBUTION OF WHISPER SCORES

Under the conditions of the method, the child who whispered in each of ten five-minute intervals would have a score of ten, while the child who did not whisper at all would have a score of zero. Scores for the elementary grades (Table II) varied between zero and eight, with a mean manifestation of 2.2. Since the score for each child in the elementary school is the average of records from two observers, the variabilities of the scores are somewhat curtailed. An examination of the table reveals the presence of a large number of zero values. In a sense, the method fails to differentiate about 33 per cent of the elementary grade children. Previous experience with the method in the measurement of other types of behavior suggests that added observations are needed in order to further differentiate the zero scores. The scores for the single observer of the twelfth grade vary between zero and ten with a mean of 4.5 and relatively few zero scores.

No reliable sex differences appear in the table. Although the data were sexed for purposes of analysis, the lack of any constant trend has led the writer to deal with only the totals in the remainder of this report.

RELATION OF WHISPER SCORES TO GRADE IN SCHOOL

No clear-cut relationship between the mean amount of whispering and the grade in school can be demonstrated with the number of cases available (Table III). The data suggest a larger amount in the first grade than in subsequent grades in the elementary school. On the other hand, a previous table (II) shows an almost equal amount in a

group of high school seniors. It would appear that the gross amount of whispering in a particular grade is a function of the situation rather than the reflection of any fundamental developmental continuum. Deviation from the mean of a given situation is probably of more significance than the raw score. Such fluctuations as are found from room to room cannot be attributed solely to errors of measurement for two observers tend to give similer descriptions (Table IV).

TABLE II.—DISTRIBUTION OF WHISPER SCORES BY SEX

Scores	Elementary grades[1]			12th grade		
	Boys	Girls	Total	Boys	Girls	Total
10	.....	.....	.....	....	2	2
9	.....	.....	.....	....	2	2
8	1	.....	1	....	....	....
7	1	4	5	1	....	1
6	2	3	5	1	....	1
5	6	4	10	3	1	4
4	6	5	11	3	1	4
3	19	13	32	5	....	5
2	15	13	28	2	....	2
1	27	30	57	1	1	2
0	32	43	75	1	3	4
N	109	115	224	17	10	27
Md	1.8	1.5	1.6	3.9	5.0	4.1
M	2.3	2.1	2.2	4.0	5.3	4.5
SD_{dis}	1.8	1.9	1.8	1.7	4.2	3.0

[1] Cases without age records omitted from this tabulation.

WHISPERING, INTELLIGENCE, AND SCHOOL ACHIEVEMENT

Intelligence quotients, based on a single group test, were available for children in the 6A grade. The coefficient of correlation between whisper scores and intelligence quotients was $-.34 \pm .11$. Intelligence quotients, based on the mean of five group tests, and honor-point ratios, based on all marks for three years of high school work, were available for the small group of high school seniors. The intercorrelations are presented in Table V. The relationship of whispering to honor-point ratios ($-.30 \pm .12$) is higher than that with intelligence quotients ($-.17 \pm .13$). Data on larger groups are necessary to establish these trends with certainty. The evidence available suggests

that whispering is not simply an adventitious circumstance, but is a reflection in part of adjustment to the classroom situation. The more intelligent and scholarly, interested in the work of the class, adjust to it and whisper less. It is conceivable that the relationship would change under varying conditions of work and interest.

Whispering and Marks in Conduct

In a sense, the validity of the whisper scores is intrinsic and reliability and validity coefficients become synonymous for the narrowly defined behavior which is the subject of study. If now a claim is made that the whisper scores are symptomatic of broader categories of behavior, external criteria must be used to evaluate the method.

Table III.—Whisper Scores by Grades in an Elementary School[1]

	Grades						Total
	1*B*	2*A*	3*B*–3*A*	4*B*–4*A*	5*B*	6*B*–6*A*	
N	30	19	45	49	22	59	224
Md	4.6	1.6	0.8	1.8	0.8	1.3	1.6
M	4.8	2.5	1.4	2.0	1.2	1.8	2.2
$SD_{dis.}$	1.5	2.0	1.6	1.4	1.0	1.3	1.8

[1] Cases without age records omitted from this tabulation.

Table IV.—Comparison of Mean Scores and Variability of Two Observers in Each Classroom

Room	*N*	Mean		$SD_{dis.}$	
		Observer 1	Observer 2	Observer 1	Observer 2
1*B*	32	5.6	4.4	2.0	2.0
2*A*–3*B*	35	2.9	2.1	2.7	1.8
4*B*–4*A*	40	2.6	2.5	1.5	1.4
6*A*	30	2.6	2.5	1.9	1.9
6*B*	31	1.9	1.2	1.2	.9
3*B*–3*A*	34	1.4	1.2	1.7	.9
4*A*–5*B*	35	1.4	1.0	1.2	.7

Children in the elementary school studied were given marks in conduct by teachers in grades three and above. The marks were in terms of letter ratings, *A*, *B*, *C*, *D*, and *E*. The summary for a semester was taken as the mark for each child. There is a suggestion

that the significance of whispering as one of the criteria of conduct increases as one goes from the freedom of the early school years to the more formal procedures of the upper grades (Table VI).

TABLE V.—RELATIONSHIP BETWEEN WHISPER SCORES AND OTHER VARIABLES IN A HIGH SCHOOL CLASS (N = 24), PER CENT

	IQ	HP ave.	CA
Whisper	−.17	−.30	.18
IQ	.....	.68	−.55
HP Ave	.....	.....	−.31

TABLE VI.—THE RELATION OF WHISPER SCORES TO MARKS IN CONDUCT IN SUCCESSIVE GRADES

Grade	N	r
6*B*	29	−.59
5*A*–5*B*	30	−.36
4*B*–4*A*	36	−.08
3*A*–4*B*	30	.10

SUMMARY

A group of students in a course in the psychology of personality collected data on the occurrence of whispering in elementary and high school children by the use of a time sampling technique. Each student made ten consecutive five-minute observations. The reliability coefficients for the average of two unpracticed observers varied between .33 and .88 in the rooms studied. No significant sex differences appeared in the mean whisper scores. The amount of whispering present in the various grades appeared to be a function of the situation rather than the reflection of a developmental trend. The evidence suggested that the more intelligent and scholarly whispered somewhat less than other children. The significance of whispering as one of the criteria entering into marks of conduct increased with the grades.

REPLY TO PROFESSOR KELLEY

KARL J. HOLZINGER

University of Chicago

In the May issue of this Journal Professor Kelley has taken objection to some comments I have made[1] regarding an example taken from his book, "Crossroads in the Mind of Man."

To save print I refer the reader to the data in the papers cited above. In my March paper I tried to do the following things:

1. Show that Thorndike's CAVD intelligence test has a common factor with three other tests and determine its correlation with this factor: ($r_{tg} = .96$).

2. As an example that a "parsimonious" explanation of correlations is desirable I took four out of nine tests from some of Professor Kelley's data and found one common factor an adequate explanation. I then gave Professor Kelleys' pattern based on these four tests and five others. This pattern was very elaborate, having one common factor, three substantial group factors and several specific factors. On p. 164 of my article I say,

> The above example is but a *fragment* of Professor Kelley's work on these data and is not included by way of criticism, but merely because the numeral work was at hand. As far as these *four tests* are concerned we hold that pattern (1) is adequate. Professor Kelley employs the elaborate pattern in the above table and interprets the common factor α as "heterogeneity, maturity, sex, and race." We argue that the factors α, β, γ, δ, ϵ, ζ, etc. are insignificant in these four tests, and that whatever common factor is found may be regarded as g. (Italics added.)

Professor Kelley (*loc. cit.*) charges me with the following errors:

(*a*) A false presentation of his argument because other elements of the problem were not considered (p. 365).

(*b*) Unawareness that when more variables are added to a given set the pattern explanation becomes more involved (p. 365).

(*c*) Overlooking this point in the interpretation of CAVD (p. 365).

(*d*) Implied charge that I (or people like me) would take four tests with common g, then take four other tests with a common g^1 and from these two results say $g = g^1$ (p. 366).

---

[1] Holzinger, Karl J.: Thorndike's CAVD is full of *G*. *Journal of Educational Psychology*, March, 1931.

With regard to point (*a*) I may say that I was aware I had taken only a part of Professor Kelley's data. Evidence of such awareness is in the paragraph cited above. I was not presenting any argument of his either falsely or otherwise.

I am aware and was aware of the fact that when more variables are added to a set, the factor pattern may become more complicated. Thus, if we have an explanation for four variables x_1, x_2, x_3, and x_4, this explanation may not suffice when we add x_5. Professor Kelley notes that we may get a *group* factor between x_1 and x_5. He does not, however, explain how the addition of more variables to the original tests with no group factor, introduces several group factors *in the original form.* An explanation is possible, and would have been more to the point than the artificial data in his Table I, p. 365.

As regards the CAVD test I see no justification for the charge that I was unaware that other tests might have given a different result. Professor Kelley's remark on p. 366 is particularly misleading: "To generalize from the relationships found in four tests to relationships supposed to lie in the mind of man begs the question, for doing so involves the assumption that the four tests sample the entire mental life." I agree with this quotation as a statement of a general principle, but not if it is applied as a criticism of what I have done with the CAVD test and three others (which were all Professor Thorndike furnished me). Note that if the number "nine" be substituted in place of the number "four" in this quotation, the same criticism applies in less degree to Professor Kelley's own data. Surely the more tests we have and the more nearly they cover the areas of mental life, the better the opportunity for a complete picture of such life.

My own study showed that CAVD is full to the extent $r = .96$ with the same thing that three other so-called intelligence tests contain to a large extent. On p. 96 of his measurement of intelligence Professor Thorndike makes the following inference from the same correlations I employed: "Intellect CAVD is very much the same thing as that which is measured by representative examinations for so-called intelligence." All I have done is to show the same fact more precisely and to indicate how much CAVD is saturated with this "same thing" present in other tests of so-called intelligence.

The charge that someone might say $g = g^1$ when these two are obtained from different tests is too absent to more than recognize as such.

I may say in conclusion that I am much more interested in getting at the truth about factor analysis than in getting the better of an argument. An argument, however, may bring to light the crucial aspects of a theory a little more vividly than one sided presentation. The point here emphasized I believe to be, that factor patterns are functions of the tests we use. They are also functions of the groups we test and of many other things. Except for this point I shouldn't have written this reply because I am certain Professor Kelley and I are in substantial agreement on all points at apparent issue.

THE INFLUENCE OF JAZZ AND DIRGE MUSIC UPON SPEED AND ACCURACY OF TYPING

MILTON B. JENSEN

Western Kentucky State Teachers College

This study was made at the Training School of the Central State Teachers College, Michigan, three weeks before the close of the spring semester, 1930. Fifth, eleventh and twelfth graders (twelve boys and thirty-eight girls) in three typing classes numbering seventeen, seventeen and sixteen respectively were used as subjects. All had had thirty-seven consecutive weeks of typing instruction in the high school. In age they ranged from fifteen to twenty-two years (average, 17.88). Speed and accuracy were measured under three conditions, hereafter called states of distraction:

1. *Normal.*—This was relatively free from distractions, being typical of the regular class procedure. Passing through the corridors was infrequent during these periods and there were few of the noises common to schools in many of our population centers. Whether this latter fact means that our subjects were more susceptible to distraction than they would be under city conditions we cannot say.

2. *During the Playing of Jazz Music on an Edison Phonograph, Cabinet Model.*—A heavy steel needle was used in playing all the records, the tempo was kept normal and no changes were made in volume control. The jazz selections were: "Valencia," Perfect Record #14625, played by the Mayflower Serenaders; "At the Prom," Victor Record #38105A, played by Irving Mills and his Modernists; and "Bugel Call Blues," Victor Record #38105B, played by Jack Pettis and his Pets.

3. *During the Playing of Dirge Music.*—The selections reproduced were: "Death of Ase," Victor Record #35470B, played by the Victor Concert Orchestra; "Thais Meditation," Victor Record #6186A, played by Fritz Kreisler accompanied by Carl Lamson; and "Indian Lament," Victor Record #6186B, played by Fritz Kreisler accompanied by Vincent O'Brien.

Practice effects were controlled by testing each of the three classes separately under all three conditions on three successive days and varying the order of testing so that, when the data for the three classes were combined, equal amounts of practice would accrue to each method. Thus class "A" was tested in the order: Tuesday—normal, jazz, dirge;

Wednesday—jazz, dirge, normal; and Thursday—dirge, normal, jazz. Class "B" was tested in the order: Tuesday—jazz, dirge, normal; Wednesday—dirge, normal, jazz; and Thursday—normal, jazz, dirge. And class "C" was tested in the order: Tuesday—dirge, normal, jazz; Wednesday—normal, jazz, dirge; and Thursday—jazz, dirge, normal. Since the classes were approximately equal in size and since the testing all came at the same hour of the day, this procedure may be thought of as supplying an adequate control of practice effects.

Three separate five minute typing tests were used, all being administered in the same order to each of the classes each of the three days: Test No. 4, p. 64, "New Rational Typewriting" by R. P. Sorelle, published by the Gregg Publishing Co.; tests put out by the Royal Typewriting Co. in May, 1929 and April, 1930. Estimates of the reliabilities of these tests were secured by correlating the scores of the students in words per minute and applying the technique devised by Dr. Shen in 1924.[1] These reliabilities in the order, Sorelle, Royal 1929, Royal 1930, were: under normal conditions, .62, .74, .57 (av., .64); during the playing of jazz, .89, .55, .68 (av., .71); during the playing of dirge, .70, .85, .46 (av., .67).

It will be seen that these examinations are accurate enough to justify their use as measures of group differences. Assuming the reliability of each test equal to the average of the three tests, for each of the three states of distraction, an application of the Spearman-Brown formula gives a reliability of .84 for an averaging of the test scores under normal conditions, a reliability of .88 during the playing of jazz, and of .86 during the playing of dirge music.[2]

While the differences in these estimates of reliability are of no great statistical importance they *do* show that the accuracy of our measures is not impaired by the distractions used. What change there is, is in the direction of improved accuracy—an increase of .04 with jazz over the normal and of .02 for dirge over the normal.

Since typing results, traditionally, are measured in words per minute the effects of the distractions upon the reliabilities of the tests were determined in this unit. In determining the effects upon student performance, however, we have examined the speed of striking the keys (number of strokes) and the number of errors separately. Table I gives the average number of strokes, errors and words per minute for

[1] Shen, E.: The Reliability Coefficient of Personal Ratings. *Journal* of *Educational Psychology*, Vol. XVI, 1925, pp. 232–236.

[2] Kelley, T. L.: "Statistical Method," p. 205.

each of the three states of distraction, together with the probable errors of these means.

TABLE I.—AVERAGE NUMBER OF STROKES, ERRORS AND WORDS PER MINUTE FOR TYPING UNDER THREE STATES OF DISTRACTION

	Normal	Jazz	Dirge
Strokes.....	215.08 ± 3.70	215.89 ± 3.82	209.99 ± 3.28
Errors.....	.937 ± .049	1.212 ± .082	.907 ± .053
Words.....	33.64 ± .92	31.06 ± 1.04	32.93 ± .86

It is seen by inspection that jazz music had no appreciable influence upon the typing speed of our subjects as measured in strokes per minute. The dirge, however, brought about an average decrease from the normal of approximately five strokes per minute—2.37 per cent. This mean difference (5.09) divided by its standard error (2.41) gives 2.11. In terms of normal probability areas, the chances are nine hundred eighty-three in one thousand that the dirge as a distraction decreased the speed of striking the keys. When the number of errors per minute of typing is examined, however, a different picture is presented. The dirge, which materially reduced the speed of striking the keys, actually reduced the number of errors, leaving the end result (number of words per minute) little different from the normal. Jazz, as a distraction, markedly increased the number of errors—.56 per cent as against .44 per cent for the normal. The mean increase in errors per minute (.275) divided by its standard error (.038) gives 7.24, leaving no doubt as to the seriousness of the influence of jazz music on typing, so far as errors are concerned. The mean number of words per minute is not appreciably lower with dirge as a distraction than under normal conditions, but jazz, despite its lack of effect upon the number of strokes per minute, seriously affects the final score in words per minute, because it is conducive to error. Our subjects, on the average, typed 2.58 fewer words per minute with jazz than under normal conditions—a decrease of 7.67 per cent. The standard error of this difference is .796 so that the difference (2.58) divided by its standard error is 3.24.

In examining the relationship between speed and accuracy, the number of strokes and the number of errors were correlated (product-moment) for the three states of distraction. They were found to be: under normal conditions, $-.111 \pm .094$; with jazz, $-.191 \pm .092$; with

dirge, $-.508 \pm 080$. Had our populations been larger, more extensive examination would have been made of the ranges of ability most seriously influenced by distractions of this sort. The relation between speed of typing and susceptibility to distraction is a problem of considerable importance to both the business man and the educator. Unfortunately, our data cannot solve the problem as the writer should wish.

It is of interest that speed and error are so little related under normal conditions. A plotting of the percentages of error given in Table II shows a marked rise throughout the middle deciles as opposed to a gradual dropping off which one might be led to expect. This situation is even more marked with jazz as a distraction. With the dirge the tendency is still present but the decrease comes earlier—in keeping with the strong negative correlation between speed and error under this type of distraction. Table II gives the mean number of strokes per minute and the mean percentages of error for the deciles of the distribution under each of the three states of distraction. The decile order is the same throughout as under normal conditions, no changes being made for changed positions with jazz and dirge as distractions.

TABLE II.—MEAN STROKES PER MINUTE AND PERCENTAGES OF ERROR BY DECILES FOR THREE STATES OF DISTRACTION

Decile	Normal		Jazz		Dirge	
	Strokes per minute	Per cent error	Strokes per minute	Per cent error	Strokes per minute	Per cent error
1	149.8	.45	148.2	.59	.56.9	.54
2	179.9	.42	192.4	.44	181.8	.46
3	152.0	.46	200.3	.78	186.7	.30
4	202.2	.55	196.7	.77	196.4	.49
5	208.2	.58	208.5	.60	204.3	.69
6	215.8	.59	215.6	.59	208.7	.52
7	224.2	.68	222.2	.97	213.4	.58
8	231.8	.36	234.2	.39	228.7	.32
9	255.6	.17	247.9	.26	243.8	.35
10	281.4	.21	285.8	.35	272.1	.27

CONCLUSIONS

Music is a serious distraction to typists under the conditions employed in this study. Jazz and dirge were used because they are

extremes. Other types of music might be expected to have effects intermediate between the influences of these two. It seems reasonable that any performance calling for rapidity and skill of movement will meet interference from stimuli as intruding as the extremes of jazz and dirge employed in this study.[1] The magnitude of this interference is indicated by our results. To what extent accommodation might offset the distraction over a protracted period of time, we have no measure. The safe procedure, where speed and accuracy are demanded, is to avoid distractions of this type unless it can be demonstrated that accommodation has taken place to such an extent that performance will not be materially affected.

---

[1] The writer has made a preliminary study of the effects of the radio upon simple arithmetic computations. Data for sixteen subjects adding single digit numbers during sixteen fifteen-minute experimental periods (eight with the radio and eight without—alternated to obviate practice effects and distributed over eight days) are strictly in keeping with the results obtained from the typing classes. Speed was not materially affected but there was a decrease of eight in the percentage of accuracy—ninety-three per cent of the additions attempted without the radio were correct and only eighty-five per cent of those attempted with the radio going were correct.

A PROOF THAT THE POINT FROM WHICH THE SUM OF THE ABSOLUTE DEVIATIONS IS A MINIMUM IS THE MEDIAN

PAUL HORST

Personnel Research Department, United States Civil Service Commission

It is quite generally known that the median value of a statistical series is that value from which the sum of the absolute deviations is a minimum. Elementary handbooks on statistical method sometimes refer to this fact but since they are not concerned with derivations, they offer no mathematical proof. Both Yule[1] and Kelley[2] demonstrate the truth of this statement by showing that the absolute deviations would be greater if taken from any other point than the median.

The writer rather incidentally discovered a proof which is more clean cut than any he has seen, and it might be of some interest to others. Instead of showing that the sum of the absolute deviations would be greater if measured from any other point than the median we may show with a more general approach that the point from which the sum of the absolute deviations is a minimum is the median.

First we derive the general expression for the average deviation from any point. We have given a statistical series,

$$X_1, X_2, X_3, \ldots X_m, X_{m+1}, X_{m+2}, \ldots X_n$$

arranged in order of magnitude, where all the values up through X_m lie below any point P, and all the values above X_m lie above this point. Then the average deviation about P will be,

$$\mathrm{AD}_P = \frac{1}{N}\left[\sum_1^m (P - X_i) - \sum_{m+1}^n (P - X_i)\right]$$

$$= \frac{1}{N}\left[mP - \sum_1^m X_i - (n - m)P + \sum_{m+1}^n X_i\right]$$

$$\mathrm{AD}_P = \frac{1}{N}\left[\sum_{m+1}^n X_i - \sum_1^m X_i - (2m - n)P\right] \qquad (1)$$

which is the same as formula (8) given by Kelley.[3]

[1] Yule, G. U.: "Introduction to the Theory of Statistics." P. 144.

[2] Kelley, T. L.: "Statistical Method." P. 74.

[3] Kelley, T. L.: "Statistical Method." P. 73.

Now we wish to determine P in (1) so that AD_P will be a minimum. To do this we simply differentiate (1) with respect to P and equate the derivative to zero. Thus,

$$\frac{d(AD_P)}{dP} = \frac{-(2m - n)}{N} = 0 \tag{2}$$

Then solving for m in terms of n we have simply

$$m = \tfrac{1}{2}n \tag{3}$$

That is, half the values are below P and half above, which is, of course, the definition of the median. Hence that point from which the sum of the absolute deviations is a minimum is the median.

RELATION BETWEEN USE OF DIFFERENT PARTS OF SPEECH IN WRITTEN COMPOSITION AND MENTAL ABILITY[1]

CHARLES P. LOOMIS AND ANNA MAY MORAN

North Carolina State College of Agriculture and Engineering

Since speech is used directly or indirectly more than any other single human attribute as a criterion by which to determine individual intelligence, it would appear advantageous to ascertain the relation of speech to mental ability from every possible angle. Most intelligence tests at the present time demand either speech reactions or reactions to speech stimuli, or both. The problem undertaken in the present study is to discover the relationship between the proportion of different parts of speech in writing vocabularies of individuals and their mental ability.

I

The importance of language in human activity is stressed by most psychologists, sociologists, and cultural anthropologists. It is generally considered to be the chief vehicle by which the experience of the past and present is transmitted. Speech has made human culture possible and man a "time binding animal." To quote from Professor Cooley: "Language is the vehicle by which human relations exist and develop—all the symbols of mind, together with the means of conveying through space and preserving them in time, expression of fact, attitude and gesture, tones of voice, printing, and everything in the way of mental growth, has its existence therein. Without language the mind does not develop a true human nature, but remains in an abnormal and non-descript state neither human nor properly brutal."[2]

So closely is language related to intelligence and the processes of developing mental ability, that it is believed that if one could measure an individual's language ability he would have a fair index to that

[1] The article is a summary of a master's thesis presented at the N. C. State College, under the direction of Dr. Karl C. Garrison, Professor of Educational Psychology.

[2] Cooley, Charles Horton: "Social Organization." Charles Scribner's Sons, York, 1909, p. 62.

individual's general mental ability. "It has often been said that thought would be impossible without words; and it is true that we can hardly conceive of human thought save as formed and embodied and expressed in language. Thought and articulate speech grew up, so to say, side by side; each implies the other; they are two sides of the same phase of mental development."[1]

A great many studies have been made to determine the relation between vocabulary and speech ability and general intelligence. For instance, Terman found correlations ranging from .85 to .95 between his vocabulary test scores and the scores made on the Stanford Revision of the Binet test.[2] Lockwood[3] found the correlation between composition scores based upon the Hudeleson scale and the semester grades of freshmen to be $.77 \pm .029$. Lockwood also found a correlation between general intelligence (Otis test) and composition scores of fifty-seven boys to be $.67 \pm .05$, and of forty-three girls to be $.76 \pm .04$, and concluded that a high degree of intelligence is directly related to ability to write adequately. Sangren[4] arrived at a similar conclusion after a study of ninth grade pupils; but Boss,[5] who studied twelfth grade pupils of superior mental ability, reported a lack of correlation between intelligence and achievement in English composition.

Closer to the problem at hand is the analysis and classification of different parts of speech used at different periods of language growth.

---

[1] Titchener, E. B.: "A Beginner's Psychology." The Macmillan Company, New York, 1917.

[2] Terman, Lewis M. *et al.*: "Genetic Studies of Genius," Volume I. Stanford University Press, 1925, p. 25.

[3] Lockwood, H. R.: Correlation of Mental Maturity of One Hundred College Freshmen and Their Ability to Write English Composition. *Unpublished Master's thesis*, Department of Education, University of Chicago, 1925.

Among studies of this kind are those of Hughes, W. H.: The Relation of Intelligence to Vocabulary and Language Training. *English Journal*, Vol. XIV, Oct., 1925, and Buller, A. D.: Intelligence Tests and Achievement in English Composition. *Educational Administration and Supervision*, Vol. XIII, Jan., 1927, and Garrison, K. C.: The Relationship between Three Different Vocabulary Abilities. *Journal of Educational Research*, Vol. XXI, Number 1, Jan., 1930, pp. 43–44.

[4] Sangren, Paul V.: Intelligence Tests and the Classification of Students in Ninth-grade English. *Educational Administration and Supervision*, Vol. IX, Dec., 1923, p. 547.

[5] Boss, Mabel E.: The Relation of Performance in Mental Tests to Achievement in High School English. *Unpublished Master's thesis*, Department of Education, University of Minnesota, 1925, p. 89.

In this regard, discrepancies in the classification of certain portions of language activity have arisen because of the difficulty of assigning the part of speech to the same word or expression under different conditions. A summary of the literature on the relative number of the various parts of speech was made by C. W. Waddle, in 1918, with the following findings:[1]

1. Interjectional speech was characteristic at the beginning.
2. Nouns were used early in relatively large numbers.
3. From the first year on, the verbal element was relatively large.
4. The proportion of adjectives to adverbs was greater at younger ages.
5. Personal pronouns, relative pronouns, and subordinating and connecting words, were acquired with difficulty.

These studies, however, have been confined to the investigation of the *growth* of the use of the different parts of speech. Boyd,[2] in his study on the development of sentence structure in childhood, devotes a section to the elements in the vocabulary. He says: "Coming now to the words which are components of clauses and sentences, we may consider the relative percentages of the different parts of speech in the total words of the collected sentences of the child (counting each word each time it occurs)."

The percentages of the different parts of speech of the child studied by Boyd do not differ greatly from those of the study which follows, but they are classified differently. His comparison of the various parts of speech used by the growing child with those of male and female novelists is interesting. The percentage of relational words (conjunctions and prepositions) in adult sentences are quite higher, but the percentage of verbs and adverbs associated with them is considerably lower. Nouns and pronouns together are of about the same frequency in children and adults. Adults tend to use more nouns than pronouns, but the children use more verbs than the novelists. There was not much difference between the men and women novelists with regard to percentage of parts of speech used.

In a recent monograph, Symonds and Lee present data gathered for the purpose of analyzing the development of vocabulary usage in written composition, and to determine the changes that occur in the

---

[1] Waddle, C. W.: "An Introduction to Child Psychology." Houghton-Mifflin Company, Boston, 1918, p. 74.

[2] Boyd, William: *British Journal of Psychology*, Vol. XVII, 1927, pp. 181–191.

use of words as the child becomes more and more mature. In their conclusion they say: "It was surprising that the per cent of changes are so slight. A few basic articles, conjunctions, prepositions, pronouns, nouns, and verbs seem to be the framework of the language and are used whenever language is used."[1] One might anticipate from this conclusion that there would be little variation in the proportions of different parts of speech used by different ages and intelligence. This was found to be somewhat true in the present study, but there was enough variation about the averages to warrant further statistical treatment involving correlations.

In a study of high school pupils' vocabularies, O'Brien[2] found that the number of adjectives used by students increased steadily as the school grade increased. This was also found to be true of words or phrases pertaining to abstract qualities or ideas. He states that "it was felt that the extent to which the pupil used well chosen adjectives would provide a significant index of the precision and clearness of his thinking as well as a gauge of that phase of his vocabulary." Of course it would be hard to treat the "well chosen adjectives" quantitatively, but the increase in the number of adjectives used as higher school grades are attained might indicate that there is a relation between mental ability and proportion of adjectives used.

II

The data for the present study were obtained from the Junior division of the New Raleigh High School, Raleigh, North Carolina. Composition work of three groups, two eighth and one high seventh grade, was used. The composition papers were collected at different intervals throughout the first term of the school year 1929–1930. In order to allow the students to have as much freedom as possible in the use of their vocabularies, the subject-matter for the compositions was left to the writers. As a consequence, the compositions vary widely in their content as a statement of their analysis show, on p. 469.

In order to keep as close to the aim of spontaneity as possible, it was necessary to reject some themes which had obviously been written

---

[1] Symonds, P. M. and Baldwin Lee: Studies in Learning of Expression, Number III—Vocabulary. *Teachers College Record*, Vol. XXXI, October, 1929, p. 50.

[2] O'Brien, F. P.: The Vocabulary of High School Pupils in Written Composition. *Journal of Educational Research*, Vol. II, 1925, pp. 334–350.

by pupils who had resorted to reference books for facts in order to write the particular theme handed in. Long lists of proper or common nouns were also omitted.

As a mode of procedure the following classifications of speech were set up: (1) Nouns and pronouns; (2) verbs and verb forms; (3) adjectives and adverbs (modifiers); (4) conjunctions and prepositions (connectors); and (5) the articles *a*, *an*, and *the*. In order to make the division of the words into different categories as objective as possible, words were assigned to their division only after ascertaining how they were used in the sentence. Sidwell and Siegfried's Handbook of Grammar[1] was used as a guide.

TABLE I.—ARBITRARY ANALYSIS OF THREE HUNDRED TWENTY-SIX COMPOSITIONS WRITTEN BY EIGHTY HIGH SCHOOL STUDENTS WHO WERE ALLOWED TO CHOOSE THEIR OWN FORM AND SUBJECT-MATTER

GENERAL THEME OF COMPOSITION	NUMBER SECURED
Description—nature, topics of interest, experiences, animal life, persons, holidays, geographical localities	107
Civics—tariff, Pan-American congress, League of Nations, construction of the Government, duty as citizens, current events of interest	91
Narrative—completion of stories, conversations, Story of Evangeline	64
Exposition—"How to Make," etc., "How to Earn Money," "How to Live at Home," "Why Study Latin"	27
Letters—personal, and letters of application	20
Moral Essay—interpretation of stories with moral	17
Total	326
Average number of compositions per pupil	4.075

After the words of each student's composition had been classified, the percentages of the total number of words used by an individual classified under the different catagories of the parts of speech were found, for each individual separately. In all cases all words were classified so that the sum of the percentages of the different parts of speech used by each individual always amounted to one hundred. The table on p. 470. shows the averages and deviations of individual percentages of different parts of speech used. Articles have the highest coefficient of variability.

In order that relationships could be found between uses of different parts of speech and intelligence, certain indices of the mental abilities of

[1] Sidwell, Paul and Russell G. Siegfried: "Handbook of Grammar." Charles Scribner's Sons, New York, 1928.

the students whose compositions were analyzed had to be obtained. The *first* measure used was the Terman Group Test of Intelligence, Form A. Mental ages were computed and used throughout the study.

TABLE II.—AVERAGE AND VARIABILITY OF PERCENTAGES OF WORDS WHICH WERE CLASSIFIED UNDER FIVE DIVISIONS OF SPEECH AS TAKEN FROM THREE HUNDRED TWENTY-SIX COMPOSITIONS WRITTEN BY EIGHTY HIGH SCHOOL STUDENTS

	Arithmetic mean of per cents used by each person	Standard deviation	Coefficient of variability[1]
Nouns	35.2	2.095	5.95
Verbs	23.8	2.624	11.00
Modifiers	14.6	2.366	16.12
Connectors	16.7	2.003	11.96
Articles	9.5	2.290	24.04

[1] Coefficient of variability was found by the formula $CV = \frac{100SD}{mean}$.

The *second* measure used was a "Sentence Vocabulary Test," devised by Garrison of State College, Raleigh, and designed to cover a grade

TABLE III.—AVERAGES AND VARIABILITY OF MEASURES OF MENTAL ABILITY AND FATHER'S OCCUPATION OF EIGHTY HIGH SCHOOL STUDENTS

Measures of intelligence	Arithmetic means	Standard deviation	Coefficient variability
Mental age	13.75	1.545	11.236
Vocabulary test	47.95	6.134	12.792
Average school grades	83.10	7.807	9.395
Father's occupation	13.08	2.690	20.565

range of from third to eighth.[1] For a *third* criterion, the average grades in all thought subjects (English, history, civics, a natural science, and language) made by pupils during the first term of the school year 1929 and 1930 were obtained from the file in the principal's

[1] This test is much like the Holley Vocabulary Scale, but made more difficult. An example of the test is as follows: To *pledge* is to . . . sedge . . . cushion . . . council . . . promise.

A study involving its use may be found in Garrison, K. C.: The Relationship between Three Different Vocabulary Abilities. *Journal of Educational Research*, Jan., 1930.

office. Also for an index of home environment, the father's occupation was used and given a quantitative rating using the Barr Scale Rating of Occupational Status.[1] A table of these indices on p. 470.

It was desirable to see what part the factor of sex played in the use of different parts of speech as well as the intelligence ratings. The

TABLE IV.—COMPARISON OF THE PERCENTAGES OF DIFFERENT PARTS OF SPEECH USED BY EIGHTY HIGH SCHOOL GIRLS AND BOYS

	Number	Nouns	Verbs	Modifiers	Connectors	Articles
Boys	39	34.9	24.0	10.1	16.7	14.3
Girls	41	35.5	23.7	15.0	16.8	9.0

following table shows that there is, on the whole, very little sex difference in the use of the different parts of speech.

It may be seen by the above table that the chief difference is the boys' and girls' use of speech in modifiers and articles. The girls use

TABLE V.—COMPARISON OF MENTAL AGES, VOCABULARY TEST SCORES, AND AVERAGE SCHOOL GRADE BY SEX

	Number	Mental age	Vocabulary ability	Average school grades
Boys	39	13.78	47.5	80.6
Girls	41	13.57	48.3	85.4

4.9 per cent more modifiers than do the boys, while the girls use 5.3 per cent fewer articles than do the boys.

Table V compares the sexes by three other criteria. The mental age and vocabulary index of the boys and girls show very little differ-

[1] In this list one hundred representative occupations, each definitely and concretely described, were rated from one to one hundred by thirty judges according to the grade of intelligence which each was believed to demand. Probable error values were computed for all these one hundred occupations and the ratings distributed. In the present study those occupations that were not included in the one hundred listed in the Barr Scale were classified as the occupations most like them. The Barr Scale Ratings list is fully described in Terman's *Genetic Studies of Genius*, Vol. I, p. 66.

ence. The girls outranked the boys by 4.8 per cent in their average school grades.

By way of further analysis, mental age, vocabulary tests, average school grades and the percentages of the different parts of speech were studied and zero order coefficients were computed to show relationships among these factors. Table VI shows the various relationships found by this computation.

TABLE VI.—THE RELATION BETWEEN THE DIFFERENT PARTS OF SPEECH AND MENTAL AGE, VOCABULARY TEST SCORES, AND AVERAGE SCHOOL GRADES

Parts of speech	Mental age	Vocabulary test score	Average school grades
Articles[1]	.52 ± .06	.48 ± .06	.51 ± .06
Connectors	.20 ± .07	.17 ± .07	.14 ± .08
Verbs	−.27 ± .07	−.17 ± .07	−.22 ± .07
Modifiers	−.10 ± .08	−.05 ± .08	.08 ± .08
Nouns	.09 ± .08	−.10 ± .08	−.06 ± .08

[1] A correlation coefficient of .39 ± .07 was found to exist between English Grades and the per cent of articles used.

So far as the sample used is concerned, Table VI shows:

1. The use of the articles *a, an* and *the,* is a better index to mental ability than is the use of any other part of speech.

2. The use of connectors next to articles shows the highest positive correlation with measures of mental ability, but the coefficients are not high.

3. The use of many verbs indicates lower mental ability with low coefficient.

4. Other parts of speech do not indicate much relationship with the criteria of mental ability used in the study.

TABLE VII.—INTERCORRELATIONS BETWEEN THE DIFFERENT PARTS OF SPEECH USED IN WRITTEN COMPOSITION BY EIGHTY HIGH SCHOOL STUDENTS

Parts of speech	Articles	Connectors	Verbs	Modifiers	Nouns
Articles	.....	.38 ± .07	−.18 ± .08	.46 ± .06	.58 ± .05
Connectors	.....	.....	−.49 ± .06	−.22 ± .07	−.04 ± .08
Verbs	.....	.....	.....	−.36 ± .07	.01 ± .08
Modifiers	.....	.....	.....	.....	−.36 ± .07

In the table above may be found a study of the interrelations among the usage of different parts of speech.

The study of the intercorrelations among the different parts of speech summarized in Table VII indicates the following facts concerning compositions studied:

1. The use of a great many words of any single part of speech, resulted in a less or nearly the same number of words of another part of speech (with the exception of articles).

2. The use of articles correlates positively with the use of modifiers, nouns, and connectors, but negatively with verbs.

In order that the relationship between the children's environment, as measured by occupation, and other attributes might be studied, various correlation coefficients between measures of mental ability and the usage of the parts of speech were computed. Data are presented in Table VIII.

TABLE VIII.—RELATIONSHIP BETWEEN THE USE OF ARTICLES, CONNECTORS, MODIFIERS, AND THE MENTAL ABILITY, VOCABULARY TEST SCORES, AVERAGE SCHOOL GRADES AND FATHER'S OCCUPATION

	Articles	Connectors	Modifiers	Mental age	Vocabulary test	Average school grades	Father's occupation
Articles		.38 ± .07	.46 ± .07	.52 ± .06	.48 ± .06	.51 ± .06	.32 ± .07
Connectors		.....	.22 ± .07	.20 ± .07	.17 ± .07	.14 ± .08	.02 ± .08
Modifiers		.....	.....	−.10 ± .07	−.05 ± .08	.08 ± .08	.05 ± .08
Mental age		.....	.....	.....	.58 ± .05	.27 ± .07	.24 ± .07
Vocabulary test		.....	.....	.....	.....	.49 ± .06	.39 ± .06
Ave. school grades		.....	.....	.....	.....	.....	.04 ± .08
Father's occupation							

The data presented in Table VIII corroborates many other studies that have found high mental ages associated with high vocabulary test scores and school grades. The table also indicates the following:

1. Father's Occupation as rated by the Barr Scale Ratings of Occupational Status[1] previously referred to, correlates positively with the use of articles, which is the best measure of intelligence so far as parts of speech are concerned.

2. Father's Occupational Ratings correlate positively and significantly with mental age and vocabulary test scores. As here rated, father's occupation does not seem to influence school grades.

[1] See footnote 1, p. 471.

III

By way of conclusion it may be stated that:

1. It is difficult to explain the correlation between the measures of mental ability and the use of articles. The correlation coefficient between the percentage of articles used and mental age was .52 ± .06; between the percentage of articles and the vocabulary test score, .48 ± .06; and between the percentage of articles and average school grades, .51 ± .06. Since there is not the chance of using articles incorrectly that there is with most other parts of speech, such as general modifiers, the percentage of articles used makes a better index of mental ability. This is especially true in view of the fact that the correlation coefficients between measures of mental ability and the use of articles are in most cases higher than correlation coefficients between the different measures of mental ability themselves. The only exception to this is the correlation between vocabulary test scores and mental ages, .58 ± .05, while the coefficient between mental age and the use of articles is .52 ± .06.

2. The positive correlations between the use of connectors and mental ability would be easier to explain than that of articles with mental ability, if one assumed that the first part of a sentence had to be kept in mind while the last part was being composed. The correlation coefficient between percentage of connectors and mental age was .20 ± .07; between percentage of connectors and vocabulary test scores, .17 ± .07; and between percentage of connectors and average school grades, .14 ± .08.

3. As would be expected in any structure, the parts of which are as interdependent and constant as the parts of a sentence, an increase in the use of one part of speech will automatically be accompanied by a decrease in the use of the other parts of speech. This holds true for all parts of speech except articles. Since articles are so closely related with nouns and modifiers, positive correlations would be expected between articles and these parts of speech, and such relationships actually exist.

4. The use of articles has almost as much relationship with occupational status as has vocabulary ability as represented by the vocabulary test scores—the former relation being represented by the correlation coefficient .39 ± .06, and the latter, .32 ± .07. Children's environment and possibly heredity, as represented by the occupational ratings seems to influence the mental ability as represented by the mental age and vocabulary test scores, and at the same time their

use of articles. School grades are not significantly influenced by parent's occupation in this study.

5. A thoroughly quantitative method of measuring the correct and incorrect usages of different parts of speech, rather than the study of percentages of words classified as being certain parts of speech, would probably throw new light on the subject of the relation between use of different parts of speech and mental ability. It is hoped that such a method can be developed and applied.

FALSIFICATION OF AGE: A FACTOR IN CHILD GUIDANCE[1]

ANNA COHEN AND NATHAN ALTROWITZ

Psychiatric Social Workers, Child Guidance Department, Hebrew Sheltering Guardian Society, Pleasantville, N. Y.

The writers have recently had the good fortune to coöperate in the handling of a case involving a factor in child guidance which is often obscured; it is that of falsification of age. This apparently harmless act may result in innumerable conflicts which give rise to abnormalities in the development of a personality.

Our instance is that of a boy who was committed to the institution as unmanageable. He is an illegitimate child whose mother is considered neurotic. The boy was born in a distant city, and the putative father deserted almost at once. After moving to her present residence, the mother married a man who is a drunkard. The mother repeatedly told the boy that he must help her and his younger brother; that he was really her mainstay because of the incapacity and worthlessness of the stepfather. The boy was jealous of his brother, yet he took a serious big-brotherly interest in him. He quarreled with his stepfather, was disobedient, and wilful. He was also known to steal frequently. When he arrived at the institution, he continued to react in much the same manner for a long time. In order that he might sooner take over the responsibilities of the home, his mother had made him believe that he was two years older than she thought he was. She explained her placing him at the institution at the lower age by saying that only by doing this could she get the institution to care for him longer and so make it possible for her to provide at least for her younger son and herself.

Thus the boy was officially considered as old as his mother supposed him to be. That even this official age was incorrect will be seen later, though whether the mother herself did not know or would not reveal his correct age cannot be determined.

His maladjustment at the institution resulted in his being placed under the care of the Child Guidance Department. The workers

[1] The writers wish to express their appreciation to Dr. Leon W. Goldrich, Executive Director; Dr. Samuel Z. Orgel, Psychiatrist; Miss Julia Goldman, Head Psychiatric Worker; and Mr. Myron B. Blanchard, Psychiatric Social Worker, all of the institution referred to in this paper, for their kind suggestions concerning the article.

noticed that there was a worried and somewhat resentful look on his face; his step was heavy and his head sagged. At school, he was considered a poor student, doing little work, and coming tardy frequently. But the social aspect of the problem was even more serious. He stole, lied, quarreled, was disobedient and did not usually mingle or play with the other children. He exhibited great glee at his success in appropriating the possessions of others without being caught, and if he were found out, he did not seem to mind.

Since there was a conflict between the figures for the age as his mother had impressed it on him and the age as found in the institution's records, the workers communicated with the authorities at his birthplace. It was learned that he was even a year younger than his mother supposed, that is, three years younger than he had come to believe himself to be.

The boy was informed of his correct age, and gradually a marked change was noticed. He said that he no longer felt physically inferior to other boys as formerly. His entire mein has become more carefree; his head is held higher, his step is lighter, and he runs and plays with the others. He also feels that he is not retarded in school, since he is in the proper grade for his age. His attitude toward school has changed; he makes it a point never to be tardy, takes a keen interest in his school work and is having no difficulty with it. Improvement is also noticeable in his social adjustment. He no longer steals, nor is he disobedient, and he seldom quarrels.

This description indicates that here at least is one cause of problems that can be dealt with directly, quickly, and adequately. It emphasizes the significance of the apparently trivial matter of giving an incorrect age to attain some special end without taking into consideration the fact that its total effect on the personality may be very far reaching.

Yet how often, for example, do the parents, misleading the authorities, start the child in school before he has reached the proper age? The causes of falsification of a child's age vary. Those who are responsible for such an act are generally motivated by one or more of several definite factors.

The mother may have no time to care for the child, she may have to go out to earn a living, or she may have a younger child to take care of. On the other hand, the child may have been unwanted. He may be an illegitimate child, whose illegitimacy will not be suspected after

the age has been falsified. Once this is done, there develop problems that are difficult of solution.

First of all, the child is forced into situations that he is too young to meet. Imagine the discouraging and asocializing effect of constant defeat in athletics in competition with much bigger boys of his purported age. It halts the child's play life and tends to make him seclusive.

Second, it requires that the child live up to two divergent standards of behavior, one at home and one at school. So there develops in his mind a conflict as to which one he is expected to uphold.

Third, it raises the child from one social status to another without adequate preparation. At the time when he should be looking forward to completing his education, he is placed in the position of breadwinner. The child therefore feels deep concern over his failure to gain the training necessary to make him capable of shouldering these responsibilities.

The combined effect of all these factors, impressed upon the growing child may eventually turn him into a serious problem.

The solution of the difficulty must necessarily lie in the return of the child into the family situation at his correct age with the standards suitable for that age.

NOTE.—Since the purpose of this paper is to present the problem of age falsification *per se*, no attempt has been made to give the analytical interpretations that would be involved in completing the picture of the case.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

CONDUCTED BY FRANCES M. FOSTER

Adolescent Education, by Frederick E. Bolton. New York: The Macmillan Company, 1931. Pp. 506.

It is perhaps no exaggeration to say that there is no other field of education concerning which more scientific knowledge is needed, and less is available, than in the field of adolescent education. There has been a dearth of scientific literature produced in this field during the past fifteen years.

"Adolescent Education" by Professor Frederick E. Bolton will meet a long felt need in this particular field. The educational world apparently has not fully recovered from the loss sustained by the departure of that remarkable student, teacher and educator to whose memory Professor Bolton has dedicated this book.

The author states that the book is intended to be an analysis and inventory of the adolescent for the purpose of finding his potentialities and his needs as determined by his unfoldment. The course of study is treated only incidentally and only as a means to an end. The author tells us that business, social and economic conditions and traditions have been the educational determinants, and that the educational theorist, not trained in psychological analysis proceeds with what he conceives to be the logical unfoldment of the subjects and arranges the traditional school subjects into what he conceives to be ideal "mental discipline."

The child psychologist, instead of securing his basic information from industries and the world at large, endeavors to secure his information directly from the children. This procedure consists, perhaps first of all, of a thorough scientific analysis of the child's physical equipment, accompanied by a thorough and scientific analysis of his instinctive tendencies, his mental equipment, and his interests all of which serve as a key to the kinds of activity best suited to the proper development of the child.

Professor Bolton points out the often unobserved fact that a scientific educational procedure builds upon what *is*, in order to achieve

what is to be. This necessitates a scientific and functional knowledge of the pre-adolescent, adolescent and the post-adolescent life of the child. Professor Bolton's book consists of sixteen chapters each giving a scientific treatment of a fundamental aspect of adolescent education. The book reveals a mastery of the scientific literature in its field. Each chapter is accompanied by a carefully selected bibliography. The book concludes with a thorough and enlightening treatment of the ever present and gigantic problem of Character Education.

DENNIS C. TROTH.

Pennsylvania State College.

A CORRECTION

Editor, THE JOURNAL OF EDUCATIONAL PSYCHOLOGY.
Dear Sir:

There is an error (of transcription) in my formula in the article, "The Sigmas of Combined Distributions Calculated from Sigmas, Means, and Frequencies of Component Distributions," in the April issue of THE JOURNAL. In Formula 7, page 310, the second term of the left-hand member representing the combined mean should be squared. This error was discovered by Doctor Louis W. Gellermann. I am also indebted to Doctor Harry Helson for reporting difficulty with the formula as it was printed.

The correct form of the formula is:

$$\frac{(\sigma_1^2 + \bar{x}_1^2)N_1 + (\sigma_2^2 + \bar{x}_2^2)N_2 + \ldots (\sigma_n^2 + \bar{x}_n^2)N_n}{N_1 + N_2 + \ldots N_n} - \left(\frac{(\bar{x}_1)N_1 + (\bar{x}_2)N_2 + \ldots (\bar{x}_n)N_n}{N_1 + N_2 + \ldots N_n}\right)^2 = \sigma^2_{1,2,\ldots n} \quad (7)$$

C. R. GARVEY.

YALE UNIVERSITY.